

72162

STIC-Biotech/ChemLib

From: STIC-ILL
Sent: Wednesday, July 31, 2002 6:32 AM
To: STIC-Biotech/ChemLib
Subject: FW: RE: 10/006,163

Importance: High

Request for you.

-----Original Message-----

From: Huynh, Phuong N.
Sent: Tuesday, July 30, 2002 6:03 PM
To: STIC-ILL
Subject: RE: 10/006,163
Importance: High

1. Please search polypeptide of SE SEQ NO: 1 (open) against commercial and interference database.
2. Also please do oligopeptide search of SEQ ID NO: 1 against commercial and interference database.

Thanks,
 Neon
 Art unit 1644
 Mail 9E12
 Tel 308-4844

Edward Hart
 Technical Info. Specialist
 STIC/Biotech
 CMI 6B02 Tel: 305-9203

Searcher: _____
 Phone: _____
 Location: _____
 Date Picked Up: 7/31/02
 Date Completed: 7/31/02
 Searcher Prep/Review: _____
 Clerical: _____
 Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
 AA Sequences: ✓
 Structures: _____
 Bibliographic: _____
 Litigation: _____
 Full text: _____
 Patent Family: _____
 Other: _____

VENDOR/COST (where applic.)

STN: _____
 DIALOG: _____
 Questel/Orbit: _____
 DRLink: _____
 Lexis/Nexis: _____
 Sequence Sys.: 02
 WWW/Internet: _____
 Other (specify): _____

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:07:31 ; Search time 13.06 Seconds

(without alignments)
585,391 Million cell updates/sec

Title: US-10-006-163-1

Sequence: 1623
1 MAAPMNGQCVVTSASRGIG.....YLPSFLRPKVIITLITSKF 313Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCOTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1623	100.0	313	US-09-019-216-1	Sequence 1, App11
2	615	37.9	323	US-09-019-216-3	Sequence 3, App11
3	222.5	13.7	256	US-08-594-808B-7	Sequence 7, App11
4	222.5	13.7	315	US-08-793-035-9	Sequence 9, App11
5	222.5	13.7	315	US-08-793-035-10	Sequence 10, App11
6	214.5	13.2	246	US-09-238-481-2	Sequence 2, App11
7	196	12.1	252	US-08-822-322-8	Sequence 8, App11
8	196	12.1	252	US-08-822-322-8	Sequence 8, App11
9	186	11.5	337	US-09-466-109-8	Sequence 3, App11
10	183.5	11.3	313	US-08-440-856A-3	Sequence 3, App11
11	178.5	11.0	257	US-09-413-814-9	Sequence 9, App11
12	176.5	10.9	295	US-09-287-097-2	Sequence 2, App11
13	171	10.5	244	US-08-375-962B-13	Sequence 13, App11
14	171	10.5	244	US-08-562-114B-13	Sequence 13, App11
15	171	10.5	244	US-08-729-594A-13	Sequence 13, App11
16	171	10.5	287	US-08-901-306-2	Sequence 2, App11
17	170.5	10.5	303	US-09-002-298-1	Sequence 1, App11
18	170.5	10.5	303	US-09-239-052-2	Sequence 1, App11
19	169	10.4	262	US-09-363-189B-6	Sequence 6, App11
20	166	10.2	287	US-09-180-271-2	Sequence 2, App11
21	164.5	10.1	318	US-09-439-313-339	Sequence 339, App11
22	163	10.0	248	US-09-385-028-11	Sequence 11, App11
23	162	10.0	247	US-08-241-766-13	Sequence 13, App11
24	159.5	9.8	316	US-08-464-400-2	Sequence 2, App11
25	159.5	9.8	316	US-08-875-273A-2	Sequence 2, App11
26	159.5	9.8	316	US-09-123-386-2	Sequence 2, App11
27	159.5	9.8	316	US-09-019-216-1	Sequence 2, App11

28	158.5	9.8	384	1	US-08-457-245-5	Sequence 5, App11
29	157.5	9.7	251	3	US-08-822-322-9	Sequence 9, App11
30	157.5	9.7	251	4	US-09-466-109-9	Sequence 9, App11
31	155.5	9.6	261	4	US-09-468-738A-29	Sequence 29, App11
32	155.5	9.6	333	4	US-08-440-856A-4	Sequence 4, App11
33	153	9.4	246	6	5229279-7	Patent No. 5229279
34	152.5	9.4	335	3	US-09-002-298-6	Sequence 6, App11
35	152.5	9.4	335	3	US-09-109-205-19	Sequence 19, App11
36	147	9.1	244	1	US-08-762-129-3	Sequence 3, App11
37	146	9.0	335	1	US-09-002-298-7	Sequence 7, App11
38	144.5	8.9	271	2	US-07-637-865-2	Sequence 2, App11
39	144	8.9	186	4	US-08-858-207A-270	Sequence 270, App11
40	141.5	8.7	309	3	US-09-109-205-1	Sequence 1, App11
41	137	8.4	244	2	US-09-090-567-2	Sequence 2, App11
42	136.5	8.4	244	1	US-08-762-129-4	Sequence 4, App11
43	136	8.4	255	4	US-08-815-225-4	Sequence 4, App11
44	135	8.3	244	1	US-08-762-129-1	Sequence 1, App11
45	132	8.1	292	4	US-09-468-738A-2	Sequence 2, App11

ALIGNMENTS

RESULT 1
US-09-019-216-1
; Sequence 1, Application US/09019216
; Patent No. 5928923
; GENERAL INFORMATION:
; APPLICANT: Iai, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN SHORT CHAIN DEHYDROGENASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fastpro for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,216
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0475 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOF01
; CLONE: 356351
; US-09-019-216-1

Query Match 100.0%; Score 1623; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 2.8e-183;

Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MAAPNNGVCVVTGASRSGRIGIALQOLCKAGATVYITGRHLDLTVNVAQEAQSLGGCCVP 60
DB 1 MAAPNNGVCVVTGASRSGRIGIALQOLCKAGATVYITGRHLDLTVNVAQEAQSLGGCCVP 60
OY 61 VVCDSSQSEVETLEQVDRQOGRDLVNNVAGVOTILNTRKAKWEPASMMDDIN 120
DB 61 VVCDSSQSEVETLEQVDRQOGRDLVNNVAGVOTILNTRKAKWEPASMMDDIN 120
OY 121 NVGLRGHFCVSYGARLWVPAGQGLIIVISSPGLQYMFNVPYGVKACDKLAADCAHE 180
DB 121 NVGLRGHFCVSYGARLWVPAGQGLIIVISSPGLQYMFNVPYGVKACDKLAADCAHE 180
OY 181 LRRHGVSCVSLMPGIVOTELLENHAKKEEYLODPVLKQFKSAFSAETTELSGKCVALA 240
DB 181 LRRHGVSCVSLMPGIVOTELLENHAKKEEYLODPVLKQFKSAFSAETTELSGKCVALA 240
OY 241 TDPNLTSLSGKVLPSCDLARRYGLRDVGRPVODYLSVLSHVSGLGLASTYLPFLR 300
DB 241 TDPNLTSLSGKVLPSCDLARRYGLRDVGRPVODYLSVLSHVSGLGLASTYLPFLR 300
OY 301 VPKWIIALYTSKF 313
DB 301 VPKWIIALYTSKF 313
```

RESULT 2

US-09-019-216-3
Sequence 2, Application US/09019216
Patent No. 5928923

GENERAL INFORMATION:

APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN SHORT CHAIN DEHYDROGENASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/019,216
FILING DATE: Filed Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0475 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 323 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 2315796

US-09-019-216-3

Query Match 37.98; Score 615; DB 2; Length 323;
Best Local Similarity 42.18; Pred. No. 5e-64;
Matches 138; Conservative 58; Mismatches 112; Indels 20; Gaps 5;

```
OY 1 MAAPNNGVCVVTGASRSGRIGIALQOLCKAGATVYITGRHLDLTVNVAQEAQSLGGCCVP 60
DB 1 MAAPNNGVCVVTGASRSGRIGIALQOLCKAGATVYITGRHLDLTVNVAQEAQSLGGCCVP 60
OY 51 AQSILGQCVPVVCDSSQSEVETLEQVDRQOGRDLVNNVAGVOTILNTRKAKWEPASMMDDIN 110
DB 51 AQSILGQCVPVVCDSSQSEVETLEQVDRQOGRDLVNNVAGVOTILNTRKAKWEPASMMDDIN 110
OY 61 ITSNGKIALYVDSNMTEVKEFLPEKIDEBEGDLTVNNVYSLKATMICKTFFD 120
DB 61 ITSNGKIALYVDSNMTEVKEFLPEKIDEBEGDLTVNNVYSLKATMICKTFFD 120
OY 111 TPASMMDDINNVGLRGHFCVSYGARLWVPAGQGLIIVISSPGLQYMFNVPYGVKAC 170
DB 111 TPASMMDDINNVGLRGHFCVSYGARLWVPAGQGLIIVISSPGLQYMFNVPYGVKAC 170
OY 121 QDPSEWMDINGVGLRNHYCYSVYARMMYERRKGLIVNVGSLGLKYFNVAYGAKREAL 180
DB 121 QDPSEWMDINGVGLRNHYCYSVYARMMYERRKGLIVNVGSLGLKYFNVAYGAKREAL 180
OY 171 DKLADCAHELRHGVSCVSLMPGIVOTELLENHAKKEEYLODPVLKQFKSAFSAE 227
DB 171 DKLADCAHELRHGVSCVSLMPGIVOTELLENHAKKEEYLODPVLKQFKSAFSAE 227
OY 228 TTELSGKCVVALATDPNLTSLSGKVLPSCDLARRYGLRDVGRPV--ODYLSLSVLSHV 285
DB 228 TTELSGKCVVALATDPNLTSLSGKVLPSCDLARRYGLRDVGRPV--ODYLSLSVLSHV 285
OY 237 STEYTKALRLANDPGLKKSGLTFEDLAQKIDFSDKHGAKNPONIRIRITLLG-T 295
DB 237 STEYTKALRLANDPGLKKSGLTFEDLAQKIDFSDKHGAKNPONIRIRITLLG-T 295
OY 286 SGLGWLASTYLPFLRVPKWIILYTSKF 313
DB 286 SGLGWLASTYLPFLRVPKWIILYTSKF 313
OY 296 MGKEEVAKYIPPOIKLPKWMVIMQSVNRF 323
DB 296 MGKEEVAKYIPPOIKLPKWMVIMQSVNRF 323
```

RESULT 3

US-08-594-808B-7
Sequence 7, Application US/08594808B
Patent No. 5804423

GENERAL INFORMATION:

APPLICANT: Klaseen, Ralf
APPLICANT: Bringer-Meyer, Stephanie
APPLICANT: Salim, Hermann
TITLE OF INVENTION: MICROBIOLOGICAL METHOD OF MAKING
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Firm of Karl F. Ross, PC
STREET: 5676 Riverdale Ave.
CITY: Bronx
STATE: New York
COUNTRY: USA

ZIP: 10471-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/594,808B
FILING DATE: 07-FEB-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Myers, Jonathan E

REGISTRATION NUMBER: 26,963

REFERENCE/DOCKET NUMBER: 19893

TELECOMMUNICATION INFORMATION:

TELEPHONE: (718) 884-6600

TELEFAX: 718/601-1099

TELEX: 620428

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 236 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-594-808B-7

Query Match	13.7%	Score 222.5;	DB 1;	Length 256;
Best Local Similarity	31.1%;	Pred. No. 8.9e-18;		
Matches	66;	Conservative	32;	Mismatches 93;
			Indels	21;
			Gaps	5

```

0Y      5 SGOSEVFTLEEQVDRQOGRDLVLYNNAYAGVOTILNTRKAEWETPASKMDDINNYGL 124
Db      69 VTDDADVIDGVAALIERD-MGPIDILINN--AGIQ-----RAPLEEPSRKMDMLSTNV 120
0Y      125 RGHFFCSYTGARLWVPAGQGLIYVSS-----PGSLQYFENVPYGVGRACRKLADC 177
Db      121 NAVEFVGQAVARHMI PRRCGRITWICSVQSESLARPGI-----APYTATKGVAKNLTGKM 174
0Y      178 AHELRRHGVSVCSTLMPGIYVOTELKEHMAKEE 209
Db      175 ATDMGRHGLQINGLAPGYFATFEMTERLVADDE 206

```

RESULT 4
US-08-793-035-9

Sequence 9, Application US/08793035
Patent No. 6011201

GENERAL INFORMATION:

APPLICANT: Slabbs, Antoni R.
APPLICANT: White, Andrew
APPLICANT: Chase, Diane
APPLICANT: Elborough, Kelvin
APPLICANT: Fentem, Phillip A.
TITLE OF INVENTION: B-ketacyl ACP Reductase Genes From
TITLE OF INVENTION: Brassica Napus
NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,035
FILING DATE: 28-JUL-1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9414622.2
FILING DATE: 20-JUL-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB95/01678
FILING DATE: 17-JUL-1995

ATTORNEY/AGENT INFORMATION:
NAME: Kammerer, Patricia A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: MDT:132

TELECOMMUNICATION INFORMATION:
TELEPHONE: 713.787.1400
TELEFAX: 713.787.1440

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 315 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-793-035-9

Query Match	13.7%	Score 222.5;	DB 3;	Length 315;
Best Local Similarity	25.7%;	Pred. No. 1,3e-17;		
Matches 75;	Conservative 50;	Mismatches 108;	Indels 59;	Gaps 9;

[illegible]

RESULT 5

```
US-08-/93035-10
? Sequence 10, Application US/08793035
? Patent No. 6011201
? GENERAL INFORMATION:
? APPLICANT: Slabas, Antoni R.
? APPLICANT: White, Andrew
? APPLICANT: Chase, Dietre
? APPLICANT: Elborough, Keiran
? APPLICANT: Pentem, Phillip A.
? TITLE OF INVENTION: B-retocycl ACP Reductase Genes From
? NUMBER OF SEQUENCES: 10
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Arnold White & Durkee
? STREET: P.O. Box 4433
? CITY: Houston
? STATE: TX
? COUNTRY: US
? ZIP: 77210-4433
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/793,035
? FILING DATE: 28-JUL-1997
? CLASSIFICATION: 800
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: GB 9414622.2
? FILING DATE: 20-JUL-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: GB95/01678
? FILING DATE: 17-JUL-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Kammerer, Patricia A.
? REGISTRATION NUMBER: 29,775
? REFERENCE/DOCKET NUMBER: MOST:132
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 713.787.1400
? TELEFAX: 713.787.1440
? INFORMATION FOR SEQ ID NO.: 10:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 315 amino acids
? TYPE: amino acid
```

RESULT 8
US-09-466-109-8

```

APPLICANT: DELLAPORTA, STEPHEN L
TITLE OF INVENTION: MATERIALS AND METHODS FOR PRODUCING
TITLE OF INVENTION: PLANTS WITH SINGLE-SEX FLOWERS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVE. N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,856A
FILING DATE: 15-MAY-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A.
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 05463-20001.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1517
TELEFAX: (202) 887-0763
TELEX: 706141
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 337 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-440-856A-3

Query Match 11.5%; Score 186; DB 1; Length 337;
Best Local Similarity 27.4%; Pred. No. 3e-13;
Matches 82; Conservative 32; Mismatches 123; Indels 62; Gaps 9;

QY 1 MAAPRNGCVYCTVGTGSRIGRGIAQLCKAGATVYI-----TGRHDLTIRVVAQDAQSL 54
Db 49 MPKRDLGGVAVYITGGARIGIGIIVLFLFKHGRRVYIADIDDAAGEL-----ASRL 99
OY 55 GGQCPVYVCDSSQSESEVFTLFEQVYREDOQGRLDVYVNNAYAGVQTLINTRNKAFMETPAS 114
Db 100 GPQVSFVACDVSEVDVYRAVDMLSRBGRLDYVCNN--AGVLGRQTPAASISLFDPA 157
OY 115 MMDINNYGLNGHYPCSVYVYAGRLVYPAGOGILYIVISSPGS-LQYFNVPYGVKAACDKL 173
Db 158 EFDRLVRYNALGAALGMKHAARAPRRAGSIVSVASVAAVLGGRPHAYTASKNHIVGL 217
OY 174 AADCAHELRRLRG--VSCYS-----LW-----PGIVQ 197
Db 218 TKNAACELRAGHVRVNCVSPGVATPMLINAWROGHDATAADARDLIDLDTVTPSDOE 277
OY 198 TELKEHNAKEEVLDDPVLKOKSA-----FSSAETTELSCGCVY---ALATDPNLTSL 248
Db 278 VEKMEVYRGATLKGPTLRPRDIAEAVLFLASDERARYISGHNLVYDGGVTTSRNLIGL 336

RESULT 10
US-09-413-814-9
Sequence 9, Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Bayer, Stefan
APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M

```

Query Match	11.0%	Score 178.5;	DB 4;	Length 257;
Best Local Similarity	28.3%;	Pred. No. 1.4e-12;		
Matches 71;	Conservative 36;	Mismatches 105;	Indels 39;	Gaps 11.
QY	5	MNGQCVWTCASHGICRGIALQLCCKAGATYI---TGRHLDITRVVAQE-AQSLGSGCQVP	60	

[illegible]

Qy 125 RGHYCSYVGARL-----MYPAGGLIVYISSPSLOQYMFNV-PYGVKAACDKLAAD 176
Db 123 QLVNCGAV---RLVNLALPHMIKRNKGQIINISSIGVLANARFSAKYVASKAALDAFSGC 179
Qy 177 CAHELRRRGVCSCVSLMPSIVOTELLKEHMAKEEVLDDPYLKOFKSAFSSAETTELSGKCV 236
Db 180 LSAEVLKHKRISTISYMPILVTRTPM-----LAPRIKYKVPPLSPEDADLIVAI 229
Qy 237 V-----ALATDPN--ILSLSGKVLPSCDLA 259
Db 230 VKRPRIATHLGRLASITYAIAPDINNIMLSIGFNLFPSTAA 272

RESULT 13

US-08-375-962B-13
Sequence 13, Application US/08375962B
Patent No. 5731195
GENERAL INFORMATION:
APPLICANT: SIMON, ANDRAS; HELLMAN, ULF; WERNSTEDT,
APPLICANT: CHRISTER, ERIKSSON, ULF
TITLE OF INVENTION: Isolated Nucleic Acid Molecule
TITLE OF INVENTION: Which Codes for A 32 Kda Protein Having 11-CIS Retinol
TITLE OF INVENTION: Dehydrogenase Activity, and Which Associates with P63,
TITLE OF INVENTION: A Portion of a Retinol Binding Protein Receptor
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect (ASCII standard)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08375,962B
FILING DATE: 20-January-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/258,418
FILING DATE: 6-October-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pasqualini, Patricia A.
REGISTRATION NUMBER: 34,894
REFERENCE/DOCKET NUMBER: LUD 5372.1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 244 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: E.coli 3-oxoacyl[acyl carrier protein]
NAME/KEY: reductase (FABG)
US-08-375-962B-13

Query Match 10.5%; Score 171; DB 1; Length 244;
Best Local Similarity 24.4%; Pred. No. 1e-11;
Matches 60; Conservative 45; Mismatches 111; Indels 30; Gaps 7;
Qy 7 GQCVVYTGASRGIGRIALQLCKAGATVYITGRHDLTLVNAQEAASLGQCVPVVCDSS 66
Db 5 GKIALVYTGASRGIGRIATTLAARGKVGIGTATSENGAQAII---SDYLGNKGLMLNWT 61
Qy 67 QSESVTLFEQVDREOQGRDLVYVNAVAGVOTILNTRKKAFTWETPASMDDINNGLRG 126

Db 62 DPASIESVLEKI-RAEFGEVDILVNN--AGI-----TRDNLMLPMKDEEMNDIIFENLSS 113
Qy 127 HIFCSYVGARLWVPAGGLIVYISS-PSLOQYMFNVPIYGVGAACDKLAADCAHELRHNG 185
Db 114 VRLSKAVNARAAKKRKHGIIITIGSVGTGNGGQANVAAARAGLIGFSKSLAREVARSNG 173
Qy 186 VSCVSLMPSIVOTEL---LKEHMAKEEVLDDPY-----LKOFKSAFSSAE 227
Db 174 ITVNVVAPGFIEITDMRALSDQDAGILNAOVPAGRLGAGOEIANNVAFILASDEAAVITGE 233
Qy 228 TTELSG 233
Db 234 TLHVNG 239

RESULT 14

US-08-562-114B-13
Sequence 13, Application US/08562114B
Patent No. 5972646
GENERAL INFORMATION:
APPLICANT: ERIKSSON ET AL.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE WHICH CODES FOR A
TITLE OF INVENTION: 32 KDA PROTEIN HAVING 11-CIS RETINOL HYDROGENASE
TITLE OF INVENTION: ACTIVITY, AND WHICH ASSOCIATES WITH P63, A
TITLE OF INVENTION: PORTION OF A RETINOL BINDING PROTEIN RECEPTOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect 5.1 and ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/562,114B
FILING DATE: 22-No. 5972646ember-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/375,962
FILING DATE: 20-January-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kohli, Vinodet
REGISTRATION NUMBER: 37,003
REFERENCE/DOCKET NUMBER: LUD 5372.2 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 244 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: E.coli 3-oxoacyl[acyl carrier protein]reductase
NAME/KEY:
US-08-562-114B-13

Query Match 10.5%; Score 171; DB 2; Length 244;
Best Local Similarity 24.4%; Pred. No. 1e-11;
Matches 60; Conservative 45; Mismatches 111; Indels 30; Gaps 7;
Qy 7 GQCVVYTGASRGIGRIALQLCKAGATVYITGRHDLTLVNAQEAASLGQCVPVVCDSS 66
Db 5 GKIALVYTGASRGIGRIATTLAARGKVGIGTATSENGAQAII---SDYLGNKGLMLNWT 61
Qy 67 QSESVTLFEQVDREOQGRDLVYVNAVAGVOTILNTRKKAFTWETPASMDDINNGLRG 126

Db 62 DPASIESVLEKI-RAEFGEVDILVNN-AGI-----TRDNLRMKDEEMNDIETNLS 113
QY 127 HYECVYTGARLWVPAGQGLIVYSS-PGSLQYFENVPYGVKAACDKLAADCAHELRRLG 185
Db 114 VFRLSKAVYRAAMKKRHRIITIGSVGTMGNGGQANVAAAKRGLGFSKSLAREVASRG 173
QY 186 VSCVSLMPCIVOTEL---LKEHMAKEEYQDPY-----LKQKSAFSSAE 227
Db 174 ITVNVVAPGFIEDMTALSDQDQAGILQAVPAGRLGAGOEIANNVAVFLASDEAAVITGE 233
QY 228 TELSG 233
Db 234 TLHVG 239

RESULT 15
US-08-729-594A-13
; Sequence 13, Application US/08729594A
; Patent No. 6280997
; GENERAL INFORMATION:
; APPLICANT: Eriksson, Ulf; Simon, Andras; Romert, Anna
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE WHICH
; TITLE OF INVENTION: CODES FOR A 32 KDA PROTEIN HAVING 11-CIS RETINOL DEHYDROGENASE
; TITLE OF INVENTION: ACTIVITY, AND WHICH ASSOCIATES WITH P63, A PORTION OF A
; TITLE OF INVENTION: RETINOL BINDING PROTEIN RECEPTOR
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,594A
; FILING DATE: 11-October-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/562,114
; FILING DATE: 22-No. 6280997ember-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/375,962
; FILING DATE: 20-January-1995
; APPLICATION DATA:
; APPLICATION NUMBER: 08/258,418
; FILING DATE: 10-June-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6280997man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5372.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: E.coli 3-oxoacyl[acyl carrier protein]reductase
; US-08-729-594A-13

QY 7 GQVCVYTGASRGIGRIALQLCRAGATVYITGRHLDTLRVVAQEOSLGQCVPVVCDSS 66
Db 5 GR1ALVTGASRGIGRIALQTLAARGKVGIGTATSENGAOAI---SDYLGANGKGLMLNVT 61
QY 67 QSEVETLEQVDRQOGRLDVYNNVAGVOTILNTFRKAFWEYTPASWDDIINNVLG 126
Db 62 DPASIESVLEKI-RAEFGEVDILVNN-AGI-----TRDNLRMKDEEMNDIETNLS 113
QY 127 HYECVYTGARLWVPAGQGLIVYSS-PGSLQYFENVPYGVKAACDKLAADCAHELRRLG 185
Db 114 VFRLSKAVYRAAMKKRHRIITIGSVGTMGNGGQANVAAAKRGLGFSKSLAREVASRG 173
QY 186 VSCVSLMPCIVOTEL---LKEHMAKEEYQDPY-----LKQKSAFSSAE 227
Db 174 ITVNVVAPGFIEDMTALSDQDQAGILQAVPAGRLGAGOEIANNVAVFLASDEAAVITGE 233
QY 228 TELSG 233
Db 234 TLHVG 239

Search completed: July 31, 2002, 15:09:36
Job time: 125 sec

Query Match 10.5%; Score 171; DB 4; Length 244;
Best Local Similarity 24.4%; Pred. No. 1e-11;
Matches 60; Conservative 45; Mismatches 111; Indels 30; Gaps 7;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 31, 2002, 15:12:06 ; Search time 17.45 Seconds

(without alignments)
1723.551 Million cell updates/sec

Title: US-10-006-163-1

Perfect score: 313
Sequence: 1 MAAPMNGVCVVTGASRGIG.....YLPSELRVPMKIALITYSKF 313

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	5.1	323	2 T32125	hypothetical prote
2	11	3.5	251	2 AH2042	3-oxoacyl-[acyl-ca
3	10	3.2	244	1 B42147	3-oxoacyl-[acyl-ca
4	10	3.2	244	2 G90812	3-oxoacyl-[acyl-ca
5	10	3.2	244	2 T12051	3-oxoacyl-[acyl-ca
6	10	3.2	244	2 T44434	3-oxoacyl-[acyl-ca
7	10	3.2	244	2 C85672	3-oxoacyl-[acyl-ca
8	10	3.2	244	2 AB0195	3-oxoacyl-[acyl-ca
9	10	3.2	244	2 AD0642	3-oxoacyl-[acyl-ca
10	10	3.2	245	2 B83462	probable short-cha
11	10	3.2	246	2 A69621	3-oxoacyl-[acyl-ca
12	10	3.2	246	2 C83961	3-oxoacyl-[acyl-ca
13	10	3.2	246	2 B89896	3-oxoacyl-[acyl-ca
14	10	3.2	248	2 G83253	probable short-cha
15	10	3.2	249	1 S05398	granaticidin polyket
16	10	3.2	254	2 G75333	3-oxoacyl-[acyl-ca
17	10	3.2	252	2 E70604	hypothetical prote
18	10	3.2	271	1 S34678	short-chain alcoh
19	10	3.2	286	2 G83378	probable short-cha
20	10	3.2	293	2 AE1927	3-ketoreductase
21	10	3.2	302	2 T44578	C-5 ketoreductase
22	10	3.2	335	2 T29604	hypothetical prote
23	10	3.2	339	2 AE2212	hypothetical prote
24	10	3.2	521	2 H75014	hypothetical prote
25	9	2.9	240	2 C83133	3-oxoacyl-[acyl-ca
26	9	2.9	241	2 C95901	probable short-cha
27	9	2.9	244	2 AH3181	short chain dehydr
28	9	2.9	245	2 B97491	hypothetical prote
29	9	2.9	246	2 B97491	hypothetical prote

30	9	2.9	247	2 S77280	3-oxoacyl-[acyl-ca
31	9	2.9	247	2 F82776	3-oxoacyl-[acyl-ca
32	9	2.9	247	2 T12020	3-oxoacyl-[acyl-ca
33	9	2.9	248	2 F82128	3-oxoacyl-[acyl-ca
34	9	2.9	248	2 F81971	probable 3-oxoacyl
35	9	2.9	248	2 E81026	3-oxoacyl-[acyl-ca
36	9	2.9	251	2 A12708	NAD/NADP dependent
37	9	2.9	255	2 AC3438	ribitol-5-phosphat
38	9	2.9	256	1 A57149	glucosyl-5-dehydr
39	9	2.9	257	2 AE3190	3-oxoacyl-[acyl-ca
40	9	2.9	266	2 F83127	probable short-cha
41	9	2.9	274	2 G83284	probable short-cha
42	9	2.9	285	2 G87298	short chain dehydr
43	9	2.9	320	2 S22450	3-oxoacyl-[acyl-ca
44	8	2.6	173	2 AD3517	probable carbonyl
45	8	2.6	177	2 I39709	fixr homology - Agr

ALIGNMENTS

RESULT 1
T32125
hypothetical protein F59E11.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32125
R:Bradshaw, H.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid F59E11.
A:Reference number: 221124
A:Accession: T32125
A:Status: preliminary; translated from GB/EMBL/DDBI
A:Molecule type: DNA
A:Residues: 1-323

A:Cross-references: EMBL:AF016685; PIDN:AA66216.1; GSPDB:GND0023; CESP:F59E11.2
A:Experimental source: strain Bristol N2; clone F59E11
C:Genetics:
A:Gene: CESP:F59E11.2
A:Map position: 5
A:Introns: 20/1; 60/1; 108/3; 135/2; 166/2; 192/3; 227/3; 280/1

Query Match 5.1%; Score 16; DB 2; Length 323;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VTGASRGIGRCIALQL 27
|||
Db 12 VTGASRGIGRCIALQL 27

RESULT 2
AH2042
3-oxoacyl-[acyl-carrier protein] reductase [imported] - Anabaena sp. (strain PCC 7120)
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AH2042
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriti, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabai, DNA Res. 6: 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2042
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-251

A:Cross-references: GB:BA000019; PIDN:BA673593.1; PID:G17130984; GSPDB:GND00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: fabg
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 3.5%; Score 11; DB 2; Length 251;
 Best Local Similarity 100.0%; Pred. No. 0.0026;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 VTGASRGIGR 21
 |||||||
 DB 14 VTGASRGIGR 24

RESULT 3
 3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) - Escherichia coli
 B42147
 N:Alternate names: 3-Ketoacyl-ACP reductase
 C:Species: Escherichia coli
 C>Date: 30-Sep-1993 #sequence_revision 31-Oct-1997 #text_change 11-Jun-1999
 C:Accession: B64853; B42147; C41856
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617
 A:Accession: B64853
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-244 <BLAT>
 A:Cross-references: GB:AE000210; GB:U00096; NID:91787332; PIDN:AMC74177.1; PID:91787335;
 A:Experimental source: strain K-12, substrain MG1655
 R:Rawlings, M.; Cronan Jr., J.E.
 J. Biol. Chem. 267, 5751-5754, 1992
 A:Title: The gene encoding Escherichia coli acyl carrier protein lies within a cluster
 A:Reference number: A42147; MUID:92210530
 A:Accession: B42147
 A:Molecule type: DNA
 A:Residues: 1-29, 'G', 31-244 <RAW>
 A:Cross-references: GB:M84991; NID:9145879; PIDN:AA23739.1; PID:9145881
 R:Verwoert, I.I.; Verbrue, E.C.; van der Linden, K.H.; Nijkamp, H.J.; Stultje, A.R.
 J. Bacteriol. 174, 2851-2857, 1992
 A:Title: Cloning, nucleotide sequence, and expression of the Escherichia coli fabD gene,
 A:Reference number: A41856; MUID:92234941
 A:Accession: C41856
 A:Molecule type: DNA
 A:Residues: 1-42, 'R', 44-45 <VER>
 A:Cross-references: GB:M87040; NID:9145885; PIDN:AA2373.1; PID:9145888
 A:Note: sequence inconsistent with the nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:97135, NCBI:P.97150)
 C:Genetics:
 A:Gene: fabG
 A:Map position: 24 min
 C:Function:
 A:Pathway: fatty acid biosynthesis
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 C:Keywords: fatty acid biosynthesis; NADP: oxidoreductase
 F:6-182/Domain: short-chain alcohol dehydrogenase homology <SAND>
 F:6-36/Region: beta-alpha-beta NADP nucleotide-binding fold
 F:151/Active site: Tyr #status predicted

Query Match 3.2%; Score 10; DB 1; Length 244;
 Best Local Similarity 100.0%; Pred. No. 0.029;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIGR 21
 |||||||
 DB 10 VTGASRGIGR 19

RESULT 4
 3-oxoacyl-[acyl-carrier-protein] reductase (Imported) - Escherichia coli (strain 0157:H7
 G90812
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: G90812
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
 gasawara, N.; Yasunaga, T.; Kunihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and 9
 A:Reference number: A96629; MUID:21156231; PMID:11258796
 A:Accession: G90812
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-244 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA834894.1; PID:913360935; GSPDB:GN00154
 A:Experimental source: strain 0157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: EC81471
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 3.2%; Score 10; DB 2; Length 244;
 Best Local Similarity 100.0%; Pred. No. 0.029;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIGR 21
 |||||||
 DB 10 VTGASRGIGR 19

RESULT 5
 3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) - Vibrio harveyi
 T12051
 C:Species: Vibrio harveyi
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: T12051
 R:Shen, Z.; Byers, D.M.
 J. Bacteriol. 178, 571-573, 1996
 A:Title: Isolation of Vibrio harveyi acyl carrier protein and the fabG, acpP and fabF
 A:Reference number: Z17396; MUID:96134997
 A:Accession: T12051
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-244 <SHS>
 A:Cross-references: EMBL:U39441; NID:91173839; PIDN:AMC3589.1; PID:91173841
 A:Experimental source: strain B392
 C:Genetics:
 A:Gene: fabG
 C:Function:
 A:Pathway: fatty acid biosynthesis
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 C:Keywords: fatty acid biosynthesis; NADP: oxidoreductase
 F:6-182/Domain: short-chain alcohol dehydrogenase homology <SAND>

Query Match 3.2%; Score 10; DB 2; Length 244;
 Best Local Similarity 100.0%; Pred. No. 0.029;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIGR 21
 |||||||
 DB 10 VTGASRGIGR 19

RESULT 6
 3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) (Imported) - Moritella maris
 T44434
 C:Species: Moritella maris
 C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 04-Mar-2000
 C:Accession: T44434
 R:Morita, N.; Ueno, A.; Tanaka, M.; Ohguya, S.; Hoshino, T.; Kawasaki, K.; Yumoto, I.
 Biotechnol. Lett. 21, 641-646, 1999
 A:Title: Cloning and sequencing of clustered genes involved in fatty acid biosynthesis
 A:Reference number: Z22768
 A:Accession: T44434
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-244 <MOR>
A:Cross-references: EMBL:AB021978; PIDN:BA85256.1
A:Experimental source: ATCC 15381
A:Genetics:
A:Note: fabg
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
C:Keywords: oxidoreductase

Query Match 3.2%; Score 10; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 12 VTGASRGIGR 21
DB 10 VTGASRGIGR 19

RESULT 7
C85672
3-oxoacyl-[acyl-carrier-protein] reductase [imported] - Escherichia coli (strain O157:H7
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: C85672
R:Perera, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimlantia, E.; Potamoussis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB8480; MUID:21074935; PMID:11206551
A:Accession: C85672
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-244 <STO>
A:Cross-references: GB:AE005174; NID:g12514637; PIDN:AA655839.1; GSPDB:GN00145; UWGP:217
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: fabg
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 3.2%; Score 10; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 12 VTGASRGIGR 21
DB 10 VTGASRGIGR 19

RESULT 8
AB0195
3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) [imported] - Yersinia pestis
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C:Accession: AB0195
R:Parhill, J.; Wren, B.W.; Thomson, N.R.; Titchell, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0195
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-244 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC90421.1; PID:g15979637; GSPDB:GN00175
C:Genetics:
A:Gene: fabg
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
C:Keywords: oxidoreductase

Query Match 3.2%; Score 10; DB 2; Length 244;

Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 12 VTGASRGIGR 21
DB 10 VTGASRGIGR 19

RESULT 9
AD0642
3-oxoacyl-[acyl-carrier protein] reductase [imported] - Salmonella enterica subsp. e
C:Species: Salmonella enterica subsp. enterica serovar Typh
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AD0642
R:Parhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churc
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Far
s, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica s
A:Reference number: AB0502; PMID:11677608
A:Accession: AD0642
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-244 <PAR>
A:Cross-references: GB:AL51382; PIDN:CAD08319.1; PID:g16502365; GSPDB:GN00176
C:Genetics:
A:Gene: SRY1234
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 3.2%; Score 10; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIGR 21
DB 10 VTGASRGIGR 19

RESULT 10
B83462
Probable, short-chain dehydrogenase PA1470 [imported] - Pseudomonas aeruginosa (strai
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 03-Aug-2001
C:Accession: B83462
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.;
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic f
A:Reference number: AB2950; MUID:20437337
A:Accession: B83462
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-245 <STO>
A:Cross-references: GB:AE004576; GB:AE004091; NID:g9947415; PIDN:AA604859.1; GSPDB:C
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1470
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 3.2%; Score 10; DB 2; Length 245;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIGR 21
DB 10 VTGASRGIGR 19

RESULT 11

A69621
 3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) [validated] - *Bacillus subtilis*
 N:Alternate names: 3-ketoacyl-acyl carrier protein reductase
 C:Species: *Bacillus subtilis*
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
 C:Accession: A69621; PC4176; T46633
 R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.: Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Ch
 A: Eritlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, C.R.; Hensut, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Huijo, M.F.
 lech, J.; Harwood, C.R.; Hensut, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Huijo, M.F.
 Koettler, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinio
 A:Authors: Lauder, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadleir, Y.; Sato, T.; Scanlon
 A:Authors: Schleich, S.; Schoeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
 Akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumelein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MUID:98044033
 A:Accession: A69621
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-246 <KUN>
 A:Cross-references: GB:Z99112; GB:AL009126; NID:92633902; PIDN:CAB13464.1; PID:92633963
 A:Experimental source: strain 168
 R:Oguro, A.; Kakeshita, H.; Takamatsu, H.; Nakamura, K.; Yamane, K.
 Gene 172, 17-24, 1996
 A:Title: The effect of Srb, a homologue of the mammalian SRP receptor alpha-subunit, on
 A:Reference number: JCA819; MUID:96257247
 A:Accession: PC4176
 A:Molecule type: DNA
 A:Residues: 230-246 <OGU>
 A:Cross-references: DDBJ:D64116; NID:q1389548; PIDN:BAH10974.1; PID:q1237012
 R:Cromar, J.E.; Morbidoni, H.R.; de Mendosa, D.
 J. Bacteriol. 178, 4794-4800, 1996
 A:Title: *Bacillus subtilis* acyl carrier protein is encoded in a cluster of lipid biosynt
 A:Reference number: Z23107; MUID:96326321
 A:Accession: T46633
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-22, 'A', 24-246 <CRO>
 A:Cross-references: EMBL:U59433; NID:q1502418; PIDN:MAC4307.1; PID:q1502421
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: fabG; srb
 A:Map position: 135-145 degrees
 C:Function:
 A:Description: EC 1.1.1.100 [validated, MUID:96326321].
 A:Pathway: fatty acid biosynthesis
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 C:Keywords: fatty acid biosynthesis; NADP; oxidoreductase
 F:5-185/Domain: short-chain alcohol dehydrogenase homology <SADH>
 F:154/Active site: Tyr #status predicted

Query Match 3.2%; Score 10; DB 2; Length 246;
 Best Local Similarity 100.0%; Pred. No. 0.029;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

12 VTGASRGIGR 21
 |||||||||
 Db 9 VTGASRGIGR 18

RESULT 12
 C83961
 3-oxoacyl-(acyl-carrier protein) reductase fabG [Imported] - *Bacillus halodurans* (strain
 C:Species: *Bacillus halodurans*
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 03-Aug-2001

C:Accession: C83961
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: C83961

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-246 <STO>
 A:Cross-references: GB:AP001515; GB:BA000004; NID:q10174886; PIDN:BAH06210.1; GSPDB:G
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: fabG
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 3.2%; Score 10; DB 2; Length 246;
 Best Local Similarity 100.0%; Pred. No. 0.029;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

12 VTGASRGIGR 21
 |||||||||
 Db 9 VTGASRGIGR 18

RESULT 13

B89896
 3-oxoacyl-(acyl-carrier protein) reductase [Imported] - *Staphylococcus aureus* (strain
 C:Species: *Staphylococcus aureus*
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: B89896
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cu, L.; O
 ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: B89896
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-246 <KUN>
 A:Cross-references: GB:BA000018; PID:q13701031; PIDN:BAH42326.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: fabG
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 3.2%; Score 10; DB 2; Length 246;
 Best Local Similarity 100.0%; Pred. No. 0.029;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

12 VTGASRGIGR 21
 |||||||||
 Db 9 VTGASRGIGR 18

RESULT 14

G83253
 probable short-chain dehydrogenase PA3128 [Imported] - *Pseudomonas aeruginosa* (strain
 C:Species: *Pseudomonas aeruginosa*
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: G83253
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laribis, K.; L
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
 A:Reference number: A82950; MUID:20437337
 A:Accession: G83253
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-246 <STO>

A:Cross-references: GB:AE004737; GB:AE004091; NID:g9949239; PIDN:AA06516.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA3128
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 3.2%; Score 10; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.029;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 84 GRDVLVNNNA 93
 |||||
 DB 79 GRDVLVNNNA 88

RESULT 15
 S05398
 granaticin polyketide ketoreductase (EC 1.1.1.-) - Streptomyces violaceoruber
 N:Alternate names: 3-hydroxycyl-CoA dehydrogenase homolog 2; granaticin polyketide synt
 C:Species: Streptomyces violaceoruber
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 11-May-2000
 C:Accession: S05398; T46535
 R:Sherman, D.H.; Malpartida, F.; Bibb, M.J.; Kleser, H.M.; Bibb, M.J.; Hopwood, D.A.
 EMBO J. 8, 2717-2725, 1989
 A>Title: Structure and deduced function of the granaticin-producing polyketide synthase
 A:Reference number: S05393; MUID:90060034
 A:Accession: S05398
 A:Molecule type: DNA
 A:Residues: 1-249 <SHE>
 A:Cross-references: EMBL:X16144; NID:g47976; PIDN:CAA34262.1; PID:g47977
 R:Richnouse, K.; Bedford, D.J.; Tornus, D.; Bechtold, A.; Bibb, M.J.; Revill, W.P.; Flos
 Chem. Biol. 5, 647-659, 1998
 A>Title: The granaticin biosynthetic gene cluster of Streptomyces violaceoruber Tu22: se
 A:Reference number: Z23045; MUID:99051446
 A:Accession: T46535
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-249 <ICH>
 A:Cross-references: EMBL:AJ011500; PIDN:CAA09651.1
 A:Experimental source: strain Tu22
 C:Genetics:
 A:Gene: graIII
 A:Note: gra-orf6
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 C:Keywords: NAD; oxidoreductase
 F:9-187/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 3.2%; Score 10; DB 1; Length 249;
 Best Local Similarity 100.0%; Pred. No. 0.029;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 84 GRDVLVNNNA 93
 |||||
 DB 81 GRDVLVNNNA 90

Search completed: July 31, 2002, 15:14:17
 Job time: 131 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:13:37 ; Search time 30.19 Seconds

(without Alignment)
1793.555 Million cell updates/sec

Title: US-10-006-163-1

Sequence: 1 MAPPMGVCVVTGASNGIG.....YLPFLAVPKMIITLTKSF 313

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size: 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL.19:**
1: sp_archaea:**
2: sp_bacteria:**
3: sp_fungi:**
4: sp_human:**
5: sp_invertebrate:**
6: sp_mammal:**
7: sp_mhc:**
8: sp_organelle:**
9: sp_plant:**
10: sp_protist:**
11: sp_rodent:**
12: sp_virus:**
13: sp_vertebrate:**
14: sp_unclassified:**
15: sp_virus:**
16: sp_bacteriap:**
17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	240	76.7	313	4	Q96LJ7
2	240	76.7	313	4	Q96B59
3	151	48.2	313	4	Q96C05
4	39	12.5	313	11	Q96D18
5	39	12.5	313	11	Q96L04
6	16	5.1	319	5	Q9N538
7	16	5.1	323	5	Q16764
8	13	4.2	250	2	Q9K486
9	13	4.2	254	2	Q93471
10	10	3.2	45	2	Q47202
11	10	3.2	244	2	Q9R833
12	10	3.2	244	2	Q86853
13	10	3.2	245	16	Q913P2
14	10	3.2	246	16	Q9KA03
15	10	3.2	246	16	Q984V1
16	10	3.2	246	16	Q99GK7

17	10	3.2	246	16	Q92R11	Q92R11 rhizobium m
18	10	3.2	248	16	Q9H296	Q9H296 pseudomonas
19	10	3.2	253	2	Q92GCI	Q92GCI streptomyc
20	10	3.2	254	10	Q949M2	Q949M2 brassica na
21	10	3.2	254	16	Q9RT26	Q9RT26 delnoccoc
22	10	3.2	255	2	Q9ALP1	Q9ALP1 saccharopol
23	10	3.2	257	16	Q987C7	Q987C7 rhizobium 1
24	10	3.2	261	2	Q93158	Q93158 bacillus su
25	10	3.2	261	5	Q76752	Q76752 drosophila
26	10	3.2	262	2	Q9RKS8	Q9RKS8 streptomyc
27	10	3.2	262	16	P96841	P96841 mycobacteri
28	10	3.2	273	2	Q9F5J1	Q9F5J1 streptomyc
29	10	3.2	286	16	Q911X3	Q911X3 pseudomonas
30	10	3.2	302	2	Q9S0N7	Q9S0N7 streptomyc
31	10	3.2	315	10	Q949M3	Q949M3 brassica na
32	10	3.2	317	10	Q93X68	Q93X68 brassica na
33	10	3.2	320	10	Q93X62	Q93X62 brassica na
34	10	3.2	325	5	Q23612	Q23612 caenorhabdi
35	10	3.2	328	10	Q93X67	Q93X67 brassica na
36	10	3.2	521	16	Q9A7A9	Q9A7A9 caulobacter
37	9	2.9	95	2	Q49154	Q49154 methylobact
38	9	2.9	240	17	Q9UY54	Q9UY54 pyrococcus
39	9	2.9	241	16	Q9HWP0	Q9HWP0 pseudomonas
40	9	2.9	244	16	Q92W70	Q92W70 rhizobium m
41	9	2.9	247	16	Q9PF16	Q9PF16 xylella fas
42	9	2.9	248	16	Q9JXR1	Q9JXR1 neisseria m
43	9	2.9	248	16	Q9JW61	Q9JW61 neisseria m
44	9	2.9	253	2	Q9K3Y8	Q9K3Y8 streptomyc
45	9	2.9	254	2	Q9F8X4	Q9F8X4 streptomyc

ALIGNMENTS

RESULT 1
Q96LJ7 PRELIMINARY; PRT: 313 AA.
AC Q96LJ7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CDNA FLJ25430 FIS. CLONE T5706262.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Suzuki Y., Hata H.,
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,
RA Nagai K., Isogai T., Sugano S.;
RT "NEO human cDNA sequencing project."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK058159; BAB71694.1;
SQ SEQUENCE 313 AA: 33909 MW: 832F83FA75D931A3 CRC64;

Query Match 76.7%; Score 240; DB 4; Length 313;
Best Local Similarity 100.0%; Pred. No. 6.3e-235;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 74 LFEQVREDOGRDLVYNNAYAGVOTILTRNKAFFETPASMDDINNYGLGHYFCSY 133
DB 74 LFEQVREDOGRDLVYNNAYAGVOTILTRNKAFFETPASMDDINNYGLGHYFCSY 133
QY 134 GARLWPAAGGLIVIVSSPSIQYMFNVPYGVGAACDLAADCALHRRHGVSCVSLP 193
DB 134 GARLWPAAGGLIVIVSSPSIQYMFNVPYGVGAACDLAADCALHRRHGVSCVSLP 193
QY 194 GIYQTELEKHAKEVLDQPVAKQFKSAFSSAFETELSGKCVVATATPDNITLSGKVL 253

DB 194 GIVOTELKEHMAKEEVLQDPVLKQKSAFSSAETTELSCGVVALADPNILSLSGVYL 253
 OY 254 PSCDLARRRGLRDVGRPVQDYLSSVSHVSGLGWLAASYLPSPFLRVPKWIITALYTSKF 313
 DB 254 PSCDLARRRGLRDVGRPVQDYLSSVSHVSGLGWLAASYLPSPFLRVPKWIITALYTSKF 313

RESULT 2
 096B59 PRELIMINARY: PRT: 313 AA.
 AC 096B59: 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHEICAL 33.9 KDA PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=MELANOMA;
 RA Strausberg R.;
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC015943; AAH15943.1; -
 KM Hypothetical protein.
 SQ SEQUENCE 313 AA; 33881 MW; EPDD252724E38776 CRC64;

Query Match 76.7%; Score 240; DB 4; Length 313;
 Best Local Similarity 100.0%; Pred. No. 6.3e-235;
 Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 74 LFEVDYRQOQRLDVLVNNAGVOTILTRNKAFWETPASMMDINNVLGRHYFCVY 133
 DB 74 LFEVDYRQOQRLDVLVNNAGVOTILTRNKAFWETPASMMDINNVLGRHYFCVY 133
 OY 134 GARLWVPGGGLIYVSSPGSLQYMFNPVGVGAACDKLAADCAHELRHGVSCVSLMP 193
 DB 134 GARLWVPGGGLIYVSSPGSLQYMFNPVGVGAACDKLAADCAHELRHGVSCVSLMP 193
 OY 194 GIVOTELKEHMAKEEVLQDPVLKQKSAFSSAETTELSCGVVALADPNILSLSGVYL 253
 DB 194 GIVOTELKEHMAKEEVLQDPVLKQKSAFSSAETTELSCGVVALADPNILSLSGVYL 253
 OY 254 PSCDLARRRGLRDVGRPVQDYLSSVSHVSGLGWLAASYLPSPFLRVPKWIITALYTSKF 313
 DB 254 PSCDLARRRGLRDVGRPVQDYLSSVSHVSGLGWLAASYLPSPFLRVPKWIITALYTSKF 313

RESULT 3
 096C05 PRELIMINARY: PRT: 313 AA.
 AC 096C05: 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE SIMILAR TO RIKEN CDNA 1110029G07 GENE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREATIC ADENOCARCINOMA;
 RA Strausberg R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC014057; AAH14057.1; -
 SQ SEQUENCE 313 AA; 33925 MW; 37FA022675C4F076 CRC64;

Query Match 48.2%; Score 151; DB 4; Length 313;

Best Local Similarity 100.0%; Pred. No. 1.3e-144;
 Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 163 YGVGAACDKLAADCAHELRHGVSCVSLMPGIVOTELKEHMAKEEVLQDPVLKQKSA 222
 DB 163 YGVGAACDKLAADCAHELRHGVSCVSLMPGIVOTELKEHMAKEEVLQDPVLKQKSA 222
 OY 223 FSSAETTELSCGVVALADPNILSLSGVLPSCDLARRRGLRDVGRPVQDYLSSVYL 282
 DB 223 FSSAETTELSCGVVALADPNILSLSGVLPSCDLARRRGLRDVGRPVQDYLSSVYL 282
 OY 283 SHVSGLGWLAASYLPSPFLRVPKWIITALYTSKF 313
 DB 283 SHVSGLGWLAASYLPSPFLRVPKWIITALYTSKF 313

RESULT 4
 09D148 PRELIMINARY: PRT: 313 AA.
 AC 09D148: 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE 1110029G07RIK PROTEIN.
 GN 1110029G07RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 RX MEDLINE=21085660; Pubmed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa K., Izawa A., Fukunishi Y., Komoto H., Adachi J., Fukuda S.,
 RA Saito T., Okazaki Y., Kiyosawa H., Kondo S., Yamataka I.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischman W., Gaasterland T., Glass C., King B., Kochita H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schram L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,
 RA Blake J., Boftell D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitlaker C., Wilmberg L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT *Functional annotation of a full-length mouse cDNA collection.*;
 RL Nature 409:685-690(2001).
 DR EMBL; AK003958; BMB23093.1; -
 DR MGI; MGI:1915960; 1110029G07RIK.
 SQ SEQUENCE 313 AA; 34031 MW; 808EBCD7910EDC03 CRC64;

Query Match 12.5%; Score 39; DB 11; Length 313;
 Best Local Similarity 100.0%; Pred. No. 5.6e-31;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GGVCVTASRGIRGIALQICAGATVYITGRHDLTLR 45
 DB 7 GGVCVTASRGIRGIALQICAGATVYITGRHDLTLR 45

RESULT 5
 099L04 PRELIMINARY: PRT: 313 AA.
 AC 099L04: 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE RIKEN CDNA 1110029607 GENE.
 GN 1110029607RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MAMMARY TUMOR. WAP-TGF ALPHA MODEL. 7 MONTHS OLD, GROSS
 RA tissue.
 RC Strauberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 DR EMBL: BC003930; AAH03930.1; -
 DR MGI: MGI:1915960; 1110029607RIK.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 313 AA; 34005 MW; 60E05BD7911BDC0C CRC64;

 Query Match 12.5%; Score 39; DB 11; Length 313;
 Best Local Similarity 100.0%; Pred. No. 5.6e-31;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 7 GQCVGTGASRGIGRGIALQLCKAGATVYITGRHLDTLR 45
 Db 7 GQCVGTGASRGIGRGIALQLCKAGATVYITGRHLDTLR 45

 RESULT 6
 Q9N538 PRELIMINARY; PRT; 319 AA.
 AC Q9N538;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE HYPOTHEICAL 35.1 KDA PROTEIN.
 GN Y32H12A.3.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE-99069613; PubMed-9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Holmes A., Elliot G., Cloud J.;
 RT "The sequence of C. elegans cosmid Y32H12A.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 DR EMBL: AC006733; AAF60486.1; -
 DR HSSP: P50162; IAF1.
 DR InterPro: IPR002198; ADH_short.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
 KW Hypothetical protein; Oxidoreductase.

SQ SEQUENCE 319 AA; 35067 MW; F6B5DDAE07EE734E CRC64;

 Query Match 5.1%; Score 16; DB 5; Length 319;
 Best Local Similarity 100.0%; Pred. No. 1.2e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 12 VTGASRGIGRGIALQL 27
 Db 10 VTGASRGIGRGIALQL 25

 RESULT 7
 Q16764 PRELIMINARY; PRT; 323 AA.
 AC Q16764;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE HYPOTHEICAL 35.8 KDA PROTEIN.
 GN F59E11.2.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE-99069613; PubMed-9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Bradshaw H.;
 RT "The sequence of C. elegans cosmid F59E11.";
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 DR EMBL: AF016685; AAG24139.1; -
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 KW Hypothetical protein; Oxidoreductase.
 SQ SEQUENCE 323 AA; 35824 MW; 4FE2AD30ED39CAE3 CRC64;

 Query Match 5.1%; Score 16; DB 5; Length 323;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 12 VTGASRGIGRGIALQL 27
 Db 12 VTGASRGIGRGIALQL 27

 RESULT 8
 Q9K486 PRELIMINARY; PRT; 250 AA.
 ID Q9K486;
 AC Q9K486;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE PUTATIVE OXIDOREDUCTASE.
 GN SC7E4.34C.
 OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Seeger K.J., Harris D.;
 RL submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kleser H.M., Denapalte D., Elchner A., Cullum J.,
 RA Kinasli H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL Mol. Microbiol. 21:77-96(1996).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 DR EMBL: AL359214; CAB94622.1; -
 DR HSSP: 012634; 1YBV.
 DR InterPro: IPR002198; ADH_short.
 DR InterPro: IPR00244; Ribosomal_L9.
 DR PRINTS: PR00080; SDRFAMILY.
 DR Oxidoreductase.
 KW
 SQ SEQUENCE 250 AA; 25506 MW; 94958BC10C82D373 CRC64;

Query Match 4.2%; Score 13; DB 2; Length 250;
 Best Local Similarity 100.0%; Pred. No. 0.00011;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VTGASRGIGRGIA 24
 Db 11 VTGASRGIGRGIA 23

RESULT 9
 ID 093471 PRELIMINARY; PRT; 254 AA.
 AC 093471;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE JADW3.
 OS Streptomyces venezuelae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=54571;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISP230;
 RA Wang L., Vining L.C.;
 RT "Control of antibiotic biosynthesis and cell differentiation in
 Streptomyces venezuelae ISP5230 by JADW1, a homolog of the gamma-
 butyrolactone autoregulators Bax and AfsA."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U24659; AAL23836.1; -
 SQ SEQUENCE 254 AA; 26434 MW; 4374AEE372B752 CRC64;

Query Match 4.2%; Score 13; DB 2; Length 254;
 Best Local Similarity 100.0%; Pred. No. 0.00011;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VTGASRGIGRGIA 24
 Db 11 VTGASRGIGRGIA 23

RESULT 10
 ID 047202 PRELIMINARY; PRT; 45 AA.
 AC 047202;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MALONYL COENZYME A-ACTYL CARRIER PROTEIN (FABD) (FRAGMENT).
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K-12;
 RX MEDLINE=92234941; PubMed=1314802;
 RA Verwoert I.I.G.S., Verbree E.C., van der Linden K.H., Nijkamp H.J.,
 RA Stultje A.R.;
 RT "Cloning, nucleotide sequence, and expression of the Escherichia coli
 fabd gene, encoding malonyl coenzyme A-acyl carrier protein
 transacylase."
 RL J. Bacteriol. 174:2851-2857(1992).
 RL EMBL: M87040; AAA23743.1; -
 FT NON_TER 45 45
 SQ SEQUENCE 45 AA; 4472 MW; E3BE39DE774E4364 CRC64;

Query Match 3.2%; Score 10; DB 2; Length 45;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VTGASRGIGR 21
 Db 10 VTGASRGIGR 19

RESULT 11
 ID 09RA33 PRELIMINARY; PRT; 244 AA.
 AC 09RA33;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 3-OXOACYL-[ACYL CARRIER PROTEIN] REDUCTASE HOMOLOG.
 GN FABG.
 OS Vibrio marinus (Moritella marina).
 OC Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;
 OC Moritella.
 OX NCBI_TaxID=90736;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MP-1 (ATCC 15381);
 RA Morita N., Ueno A., Tanaka M., Ohgiya S., Hoshino T., Kawasaki K.,
 RA Yumoto I., Ishizaki K., Okuyama H.;
 RT "Cloning and sequencing of clustered genes involved in fatty acid
 biosynthesis from the dodecahexanoic acid-producing bacterium, Vibrio
 biotechnol. Lett. 21:641-646(1999).
 RL -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 CC EMBL: AB021978; BAA05256.1; -
 DR HSSP: P19992; 1HDC.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 DR Oxidoreductase.
 KW
 SQ SEQUENCE 244 AA; 25558 MW; 871DD5CF7B7F00E3 CRC64;

Query Match 3.2%; Score 10; DB 2; Length 244;

Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIGR 21
Db 10 VTGASRGIGR 19

RESULT 12

086853

ID 086853 PRELIMINARY; PRT; 244 AA.

AC 086853;

DT 01-NOV-1998 (TREMblrel. 08, Created)

DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE 3-KETOACYL-ACP/COA REDUCTASE.

GN ORFX.

OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_Taxid=1902;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-M145;

RA Takano E., Chakraborty R., Nihira T., Yamada Y., Bibb M.;

RT "Characterisation of scbr, and scda genes involved in gamma-butyrolactone binding and synthesis in Streptomyces coelicolor.";

RL Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.

CC EMBL: AJ007731; CA007629.1; -

DR HSP: Q12634; 1YBV.

DR InterPro: IPR002198; ADH_short.

DR Pfam: PF00106; adh_short; 1.

DR PRINTS: PR00080; SDRFAMILY.

DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.

KM Oxidoreductase.

SQ SEQUENCE 244 AA; 25370 MW; 1FC1FBC0B9DE660 CRC64;

Query Match 3.2%; Score 10; DB 2; Length 244;

Best Local Similarity 100.0%; Pred. No. 0.12;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 SRGIGRGIAL 25
Db 5 SRGIGRGIAL 14

RESULT 13

0913P2

ID 0913P2 PRELIMINARY; PRT; 245 AA.

AC 0913P2;

DT 01-MAR-2001 (TREMblrel. 16, Created)

DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE PROBABLE SHORT-CHAIN DEHYDROGENASE.

GN PA1470.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_Taxid=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 15692 / PA01;

RA Stever C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,

RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Galtier L., Tolentino E., Westbrock-Wadman S., Yuan Y.,

RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reizer J., Seier M.H., Hancock R.E.W., Lory S., Olson M.V.;

*Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen.";

RL Nature 406:959-964(2000).

CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES

(SDR) FAMILY.

CC EMBL: AE004576; AAC04859.1; -

DR HSP: Q12634; 1YBV.

DR InterPro: IPR002198; ADH_short.

DR Pfam: PF00106; adh_short; 1.

DR PRINTS: PR00080; SDRFAMILY.

KM Complete proteome: Oxidoreductase.

SQ SEQUENCE 245 AA; 25365 MW; 342A1CF12C500744 CRC64;

Query Match 3.2%; Score 10; DB 16; Length 245;

Best Local Similarity 100.0%; Pred. No. 0.12;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIGR 21
Db 10 VTGASRGIGR 19

RESULT 14

09KA03

ID 09KA03 PRELIMINARY; PRT; 246 AA.

AC 09KA03;

DT 01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE 3-OXOACYL-(ACYL-CARRIER PROTEIN) REDUCTASE (EC 1.1.1.100).

GN FABG OR BH2491.

OS Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_Taxid=86665;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-C-125 / JCM 9153;

RX MEDLINE-20512582; PubMed-11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,

RA Horikoshi K.;

*Complete genome sequence of the alkaliphilic bacterium Bacillus

halodurans and genomic sequence comparison with Bacillus subtilis."

RL Nucleic Acids Res. 28:4317-4331(2000).

CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES

(SDR) FAMILY.

CC EMBL: AP001515; BAB06210.1; -

DR HSP: P19992; IHDC.

DR InterPro: IPR002198; ADH_short.

DR PRINTS: PR00080; SDRFAMILY.

DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.

KM Complete proteome: Oxidoreductase.

SQ SEQUENCE 246 AA; 26126 MW; 852B95EBBDEE9E90 CRC64;

Query Match 3.2%; Score 10; DB 16; Length 246;

Best Local Similarity 100.0%; Pred. No. 0.12;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIGR 21
Db 9 VTGASRGIGR 18

RESULT 15

0984V1

ID 0984V1 PRELIMINARY; PRT; 246 AA.

AC 0984V1;

DT 01-OCT-2001 (TREMblrel. 18, Created)

DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)

DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)

DE PROBABLE OXIDOREDUCTASE.

GN ML7832.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Phyllobacteriaceae; Mesorhizobium.
 CX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MAFF303099;
 RX MEDLINE-21082930; PubMed-11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
 Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP003012; BAB54212.1; -;
 DR InterPro; IPR002198; ADH_short.
 DR Pfam; PF00106; adh_short.1.
 DR PRINTS; PR00080; SDRFAMILY.
 KW Complete proteome.
 SQ SEQUENCE 246 AA; 26177 MW; 86346F3BB71B3932 CRC64;

Query Match 3 2% Score 10; DB 16; Length 246;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 VVTGASRGIG 20
 |||||
 DB 11 VVTGASRGIG 20

Search completed: July 31, 2002, 15:17:21
 Job time: 22; sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:13:57 ; Search time 13.51 seconds

(without alignments)
897.055 Million cell updates/sec

Title: US-10-006-163-1

Perfect score: 313
Sequence: 1 MAAPMNGVCVVTGASRGIG.....VLPSFLNPKKILALYTSKF 313

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size: 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt_40.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	3.5	271	1	NORL_ASPPA 000278 aspergillus
2	10	3.2	244	1	FABG_ECOLI 025716 escherichia
3	10	3.2	244	1	FABG_SALTY 085141 salmonella
4	10	3.2	244	1	FABG_VIRBA 055336 vibrio harv
5	10	3.2	246	1	FABG_BACST 051831 bacillus su
6	10	3.2	249	1	DHR2_STEPT 016543 streptomyc
7	10	3.2	271	1	SDR1_PICAB 008632 picea abies
8	9	2.9	244	1	FABG_VIBCH 09kq77 vibrio chol
9	9	2.9	247	1	FABG_PSEAE 054438 pseudomonas
10	9	2.9	247	1	FACI_STYNY 073574 pseudocyst
11	9	2.9	256	1	GNO_GLUOX 050199 gluconobact
12	9	2.9	320	1	FABG_CUPLA 028643 cuphea lanc
13	8	2.6	263	1	FABG_YEAST 040580 saccharomyc
14	8	2.6	276	1	DHCA_HUMAN 016132 homo sapien
15	8	2.6	278	1	FIXR_BRAJA 005406 bradyrhizob
16	8	2.6	319	1	FABG_ARATH 033207 arabidopsis
17	8	2.6	345	1	DCOP_RICPR 092c83 ticketisia
18	8	2.6	482	1	CLS2_BACST 071040 bacillus su
19	8	2.6	579	1	FD21_MYCLE 054200 mycobacteri
20	7	2.2	106	1	YDFG_BACNO 039884 bacteroides
21	7	2.2	138	1	HEX9_ADE07 003283 human adeno
22	7	2.2	145	1	ZUR_BACSU 054479 bacillus su
23	7	2.2	173	1	HR12_MOUSE 052760 mus musculu
24	7	2.2	185	1	YSV6_CAEEL 010012 caenorhabdi
25	7	2.2	241	1	PHBB_ZOORA 023228 zoogloea ra
26	7	2.2	246	1	FABG_THEMA 09x248 thermotoga
27	7	2.2	248	1	PHAB_ACIPI 050203 actinobact
28	7	2.2	248	1	Y4EK_RHISN 055434 thizobium s
29	7	2.2	248	1	YDFG_SALTY 040864 salmonella
30	7	2.2	251	1	Y484_MYCTU 011150 mycobacteri
31	7	2.2	253	1	YOHF_ECOLI 033368 escherichia
32	7	2.2	254	1	IDNO_ECOLI 039345 escherichia
33	7	2.2	256	1	Y019_THEMA 056318 thermotoga

34	7	2.2	261	1	ACT3_STRCO 016544 streptomyc
35	7	2.2	261	1	DHR2_STRCM 041177 streptomyc
36	7	2.2	287	1	PTRI_LEIMA 001782 leishmania
37	7	2.2	289	1	PTRI_LEITA 042556 leishmania
38	7	2.2	292	1	E434_ADE09 089083 human adeno
39	7	2.2	294	1	E434_ADE02 003239 human adeno
40	7	2.2	314	1	LPPW_MYCTU 010823 mycobacteri
41	7	2.2	316	1	TALA_ECOLI 078258 escherichia
42	7	2.2	383	1	CYSL_SPIOL 032260 spinacia ol
43	7	2.2	427	1	PYRC_BACCL 046538 bacillus ca
44	7	2.2	459	1	RCA_LYCPN 049074 lycopersico
45	7	2.2	485	1	VST2_HEVRH 000270 hepatitis e

ALIGNMENTS

RESULT	ID	STANDARD	PRT	271 AA
1	NORL_ASPPA			
1	NORL_ASPPA			
AC	000278;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Aflatoxin biosynthesis ketoreductase NOR-1 (EC 1.1.1.-).			
CN	NOR-1 OR NOR-1			
OS	Aspergillus parasiticus.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
OX	NCBI_TaxID=5067;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRATN-ATCC 163 / NRRL 5862 / SU-1;			
RX	MEDLINE=95085270; PubMed=7993094;			
RA	Trail F., Chang P.-K., Cary J., Linz J.E.;			
RT	Structural and functional analysis of the nor-1 gene involved in the			
RT	biosynthesis of aflatoxins by Aspergillus parasiticus.;			
RL	Appl. Environ. Microbiol. 60:4078-4085(1994).			
CC	-1- FUNCTION: MAY BE INVOLVED IN THE DEHYDRATION OF NORISOLINIC ACID			
CC	TO FORM AVERANTIN.			
CC	-1- PATHWAY: AFLATOXIN BIOSYNTHESIS.			
CC	-1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES			
CC	(SDR) FAMILY. STRONG. TO E. NIDULANS STCE.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; L27801; AAA58798.1; -			
DR	InterPro: IPR002198; ADH_short.			
DR	Pfam: PF00106; adh_short.1.			
DR	PROSITE: PS00061; ADH_SHORT; FALSE_NEG.			
KW	Oxidoreductase.			
FT	NP_BIND 28 52 NAD OR NADP (BY SIMILARITY).			
FT	ACT_SITE 185 185 BY SIMILARITY.			
FT	SEQUENCE 271 AA; 29569 MW; 34059A857672059A CRC64;			
QY	12 VTGASRGIRG 22	3.5%; Score 11; DB 1; Length 271;		
DB	29 VTGASRGIRG 39	Best local Similarity 100.0%; Pred.No. 0.0014;		
		Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
RESULT	2			
	FABG_ECOLI			

ID FABG_ECOLI STANDARD: PRT: 244 AA.
 AC P25716: P78221: (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 40, Last annotation update)
 DT 16-OCT-2001 (Rel. 41, Last annotation update)
 DE 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) (3-ketoacyl-
 GN acyl carrier protein reductase).
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=92210530; PubMed=1556094;
 RA Rawlings M., Cronan J.E. Jr.;
 RT "The gene encoding Escherichia coli acyl carrier protein lies within
 RL a cluster of fatty acid biosynthetic genes.";
 RL J. Biol. Chem. 267:5751-5754(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=9742617; PubMed=9278503;
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Moromura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horikuchi T.;
 RT "A 7.8-kb DNA sequence of the Escherichia coli K-12 genome
 RL corresponding to the 12,7-28.0 min region on the linkage map.";
 RL DNA Res. 3:137-155(1996).
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) -> 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
 CC PATHWAY.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 DR EMBL: M84991; AAA23739.1;
 DR EMBL: AE000210; AAC74177.1;
 DR EMBL: D90745; BAA35901.1;
 DR PIR: B42147; B42147.
 DR HSSP: P29132; IDFI.
 DR EcGene: EGI1318; fabG.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short.1.
 DR PRINTS: PRO0080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT.1.
 KM Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
 NP BIND 10 NADP (BY SIMILARITY).
 FT ACT_SITE 151 151 BY SIMILARITY.
 FT CONFLICT 30 30 A -> G (IN REF. 1).
 SQ SEQUENCE 244 AA; 25560 MW; 48BC1F2A7E7EEFD9 CRC64;

Query Match 3.2%; Score 10; DB 1; Length 244;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 VTGASRGIGR 21
 Db 10 VTGASRGIGR 19
 RESULT 3
 ID FABG_SALTY STANDARD: PRT: 244 AA.
 AC 085141;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) (3-ketoacyl-
 DE acyl carrier protein reductase).
 GN FABG OR STM1195 OR ST11234.
 OS Salmonella typhimurium, and
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602, 601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-S. typhimurium; STRAIN-LT2;
 RX MEDLINE=98317265; PubMed=9642179;
 RA Zhang Y., Cronan J.E. Jr.;
 RT "Transcriptional analysis of essential genes of the Escherichia coli
 RL fatty acid biosynthesis gene cluster by functional replacement with
 RL the analogous Salmonella typhimurium gene cluster.";
 RL J. Bacteriol. 180:3295-3303(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-S. typhimurium; STRAIN-LT2 / SGC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grevai N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RL LT2.";
 RL Nature 413:852-856(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-S. typhi; STRAIN-CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Kung'u K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
 RA Cronin A., Davis P., Davies R.M., Doid L., White N., Farrar J.,
 RA Felwell L., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skellon J., Stevens K.,
 RA Whitehead S., Barrrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RL enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) -> 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
 CC PATHWAY.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: AF044668; AAC38650.1; -
DR EMBL: AE008752; AAL20124.1; -
DR EMBL: AL627269; CAD08319.1; -
DR HSSP: P19992; 1HDC.
DR StyGene; SG10638; fabG.
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; 1.
KW Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 10 34 NADP (BY SIMILARITY).
FT ACT_SITE 151 151 BY SIMILARITY.
SQ SEQUENCE 244 AA; 25545 MW; 86D71DA3E9AF0363 CRC64;

Query Match 3.2%; Score 10; DB 1; Length 244;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIGR 21
|||||
DB 10 VTGASRGIGR 19

RESULT 4
FABG_VIBHA STANDARD; PRT; 244 AA.
ID P55336;
AC P55336;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) (3-ketoacyl-acyl carrier protein reductase).
GN FABG.
OS Vibrio harveyi.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
OX NCBI_TaxID=669;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B392;
RA MEDLINE-96134997; PubMed-8550484;
RA Shen Z., Byers D.M.;
RT "Isolation of Vibrio harveyi acyl carrier protein and the fabG, acpP, and fabF genes involved in fatty acid biosynthesis";
RL J. Bacteriol. 178:571-573(1996).
CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] + NADP(+) -> 3-oxoacyl-[acyl-carrier protein] + NADPH.
CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: U39441; AAC43589.1; -
DR HSSP: P19992; 1HDC.
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; 1.
KW Fatty acid biosynthesis; Oxidoreductase; NADP.
FT NP_BIND 10 34 NADP (BY SIMILARITY).
FT ACT_SITE 151 151 BY SIMILARITY.

SQ SEQUENCE 244 AA; 25519 MW; FC41A1C65B8CDAAA CRC64;

Query Match 3.2%; Score 10; DB 1; Length 244;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIGR 21
|||||
DB 10 VTGASRGIGR 19

RESULT 5
FABG_BACSU STANDARD; PRT; 246 AA.
ID P51831; O31733;
AC P51831; O31733;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) (3-ketoacyl-acyl carrier protein reductase).
GN FABG.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RA MEDLINE-96326321; PubMed-8759840;
RA Morbidoni H.R., de Mendoza D., Cronan J.E. Jr.;
RT "Bacillus subtilis acyl carrier protein is encoded in a cluster of lipid biosynthesis genes";
RL J. Bacteriol. 178:4794-4800(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-172 FROM N.A.
RC STRAIN-168;
RA MEDLINE-98195738; PubMed-9534248;
RA Foulger D., Errington J.;
RT "A 28 kbp segment from the spvW region of the Bacillus subtilis 168 genome";
RL Microbiology 144:801-805(1998).
RN [4]
RP SEQUENCE OF 230-246 FROM N.A.
RC STRAIN-168;
RA Oguro A., Kakeshita H., Takamatsu H., Nakamura K., Yamane K.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] + NADP(+) -> 3-oxoacyl-[acyl-carrier protein] + NADPH.
CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: U59433; AAC44307.1; -
DR EMBL: Z99112; CAB34364.1; -
DR EMBL: Y13937; CAA474250.1; -
DR EMBL: D64116; BAA10974.1; -
DR HSSP: P50162; IAE1.
DR Subtilist; BG11535; fabG.

DR InterPro: IP0002198; ADH_short.
 DR Pfam: PF00106; adh_short.1.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; FALSE_NEG.
 KM Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
 FT NP_BIND 9 33 NADP (BY SIMILARITY).
 FT ACT_SITE 154 154 BY SIMILARITY.
 FT CONFLICT 23 23 D -> A (IN REF. 1).
 SQ SEQUENCE 246 AA; 26282 MW; C6A391167D3237DC CRC64;

Query Match 3.2%; Score 10; DB 1; Length 246;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIGR 21
 DB 9 VTGASRGIGR 18

RESULT 6
 DHK2_STRVN STANDARD: PRT; 249 AA.
 AC P16543;
 DT 01-AUG-1990 (Rel. 15; Created)
 DT 01-AUG-1990 (Rel. 15; Last sequence update)
 DT 01-OCT-1996 (Rel. 34; Last annotation update)
 DE Granaticin polyketide synthase putative ketoacyl reductase 2
 DE (EC 1.3.1.-) (ORF6).
 OS Streptomyces violaceoruber.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1935;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=TU22;
 RX MEDLINE=90060034; PubMed=2583128;
 RA Sherman D.H., Malpartida F., Bibb M.J., Kleser H.M., Bibb M.J.,
 RA Hopwood D.A.;
 RT "Structure and deduced function of the granaticin-producing
 RT polyketide synthase gene cluster of Streptomyces violaceoruber
 RT Tu22".
 RL EMO J. 8:2717-2725(1989).
 CC -1- PATHWAY: BIOSYNTHESIS OF THE POLYKETIDE ANTI-BIOTIC GRANATICIN
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X16300; CAA34367.1;
 DR EMBL: X16144; CAA34262.1;
 DR PIR: S05398; S05398.
 DR HSSP: Q12634; 1YBV.
 DR InterPro: IP0002198; ADH_short.
 DR Pfam: PF00106; adh_short.1.
 DR PROSITE: PS00061; ADH_SHORT; FALSE_NEG.
 KM Antibiotic biosynthesis; Oxidoreductase; NAD
 FT NP_BIND 12 36 NAD (BY SIMILARITY).
 FT ACT_SITE 157 157 BY SIMILARITY.
 SQ SEQUENCE 249 AA; 25977 MW; 922832B8FE34A8B CRC64;

Query Match 3.2%; Score 10; DB 1; Length 249;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 84 GRDVLVNNNA 93

DB 81 GRDVLVNNNA 90

RESULT 7
 SDR1_PICAB STANDARD: PRT; 271 AA.
 AC Q08632;
 DT 01-OCT-1996 (Rel. 34; Created)
 DT 01-OCT-1996 (Rel. 34; Last sequence update)
 DT 01-OCT-1996 (Rel. 34; Last annotation update)
 DE Short-chain type dehydrogenase/reductase (EC 1.-.-).
 OS Picea abies (Norway spruce) (Picea excelsa).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
 OX NCBI_TaxID=3329;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP MEDLINE=94120027; PubMed=8290650;
 RA Bauer S., Galliano H., Pfeiffer F., Messner B., Sandermann H.,
 RA Ernst D.;
 RT "Isolation and characterization of a cDNA clone encoding a novel
 RT short-chain alcohol dehydrogenase from Norway spruce (Picea abies L.
 RT Karst)."
 RL Plant Physiol. 103:1479-1480(1993).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X74115; CAA52213.1;
 DR HSSP: Q12634; 1YBV.
 DR InterPro: IP0002198; ADH_short.
 DR Pfam: PF00106; adh_short.1.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 KM Oxidoreductase.
 FT NP_BIND 25 49 NAD OR NADP (BY SIMILARITY).
 FT ACT_SITE 179 179 BY SIMILARITY.
 SQ SEQUENCE 271 AA; 28724 MW; 743EACB82F85FE3 CRC64;

Query Match 3.2%; Score 10; DB 1; Length 271;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIGR 21
 DB 26 VTGASRGIGR 35

RESULT 8
 FABG_VIBCH STANDARD: PRT; 244 AA.
 AC Q9K0H7;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) (3'-ketoacyl-
 DE acyl carrier protein reductase).
 GN FABG OR VC2021.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=666;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=EL TOR N16961 / SEROTYPE O1;

```

RX MEDLINE-20406833; PubMed-10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Winm M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unanue L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermlaeva M.D., Vamathevan J., Baas S., Qin H., Dragol I., Sellers P.,
RA McDonald L., Uitterbeck T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Melnikos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
CC NADP(+) -> 3-oxoacyl-[acyl-carrier protein] + NADPH.
CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
CC PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AE004276; AAF95169.1; ALT_INIT.
CC DR TIGR: VC2021;
CC DR InterPro: IPR002198; ADH_short.
CC DR Pfam: PF00106; adh_short; 1.
CC DR Pfam: PF00678; adh_short; 2.
CC DR PRINTS: PR00080; SDRFAMILY.
CC DR PROSITE: PS00061; ADH_SHORT; 1.
CC DR Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
CC NP_BIND 10 34 NADP (BY SIMILARITY).
CC FT ACT_SITE 151 151 BY SIMILARITY.
CC FT SEQUENCE 244 AA; 25566 MW; 9FB2EB278D7CC3CE CRC64;
CC SO

```

```

Query Match 2.9%; Score 9; DB 1; Length 244;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 12 VTGASRGIG 20
DB 10 VTGASRGIG 18

```

```

RESULT 9
FABG_PSEAE STANDARD; PRT; 247 AA.
ID FABG_PSEAE
AC 054438;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) (3-ketoacyl-
DE acyl carrier protein reductase).
DE GN FABG OR PA2967.
OS Pseudomonas aeruginosa.
OS Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99395062; PubMed-10464226;
RA Kishida A.J., Hoang T.T., Schweizer H.P.;
RT "Characterization of a Pseudomonas aeruginosa fatty acid biosynthetic
RT coenzyme A:ACP transacylase (fabD)."
RT J. Bacteriol. 181:5498-5504(1999).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-ATCC 15692 / PAOI;
RX MEDLINE-20437337; PubMed-10984043;
RA Stover C.K., Pham X.-O.T., Ervin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey R.J., Britkhan F., Tolentino E., Westbrock-Madman S., Yuan Y.,
RA Garner R.L., Goltry L., Tolentino E., Westbrock-Madman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lathin K., Lim R.M.,
RA Smith K.J., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
CC NADP(+) -> 3-oxoacyl-[acyl-carrier protein] + NADPH.
CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
CC PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U91631; AAB94395.1; -.
CC DR EMBL: AE004722; AAG06355.1; -.
CC DR HSP: 012634; 1BYV.
CC DR InterPro: IPR002198; ADH_short.
CC DR Pfam: PF00106; adh_short; 1.
CC DR PRINTS: PR00080; SDRFAMILY.
CC DR PROSITE: PS00061; ADH_SHORT; 1.
CC DR Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
CC NP_BIND 10 34 NADP (BY SIMILARITY).
CC FT ACT_SITE 154 154 BY SIMILARITY.
CC FT SEQUENCE 247 AA; 25585 MW; ABB3A2B95027445B CRC64;
CC SO

```

```

Query Match 2.9%; Score 9; DB 1; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 12 VTGASRGIG 20
DB 10 VTGASRGIG 18

```

```

RESULT 10
FAGL_SYNY3 STANDARD; PRT; 247 AA.
ID FAGL_SYNY3
AC P73574;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-oxoacyl-[acyl-carrier protein] reductase 1 (EC 1.1.1.100) (3-
DE ketoacyl-acyl carrier protein reductase 1).
DE GN FAGL OR SLR0886.
OS Synechocystis sp. (strain PCC 6803).
OS Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97061201; PubMed-8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuo A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions."
RL DNA Res. 3:109-136(1996).

```

CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) - 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
 CC PATHWAY.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D90907; BA117614.1; -
 DR HSSP: P50163; 2AE1.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 KW Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
 FT NP_BIND 11 35 NADP (BY SIMILARITY).
 FT ACT_SITE 156 156 BY SIMILARITY.
 FT SEQUENCE 247 AA; 25724 MW; 91EBF9409C777F20 CRC64;

Query Match 2.9%; Score 9; DB 1; Length 247;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VTGASRGIG 20
 Db 11 VTGASRGIG 19
 |||||

RESULT 11
 ID_GNO_GLUOX STANDARD; PRT; 256 AA.
 AC P50199;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucuronate 5-dehydrogenase (EC 1.1.1.69) (5-keto-D-glucuronate 5-
 DE reductase).
 GN GNO.
 OS Glucobacter oxydans (Glucobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Glucobacter.
 CC NCB1_TaxID=442;
 RN [1]
 RP SEQUENCE FROM N.A., SEQUENCE OF 3-19, AND CHARACTERIZATION.
 RC STRAIN=DSM 3503;
 RX MEDLINE=95270578; PubMed-7751271;
 RA Kjaasen R., Bringer-Meyer S., Sahm H.;
 RT Biochemical characterization and sequence analysis of the
 RT glucuronate:NADP 5-oxidoreductase gene from Glucobacter oxydans.*;
 RT J. Bacteriol. 177:2637-2643(1995).
 RL J.
 CC -1- FUNCTION: INVOLVED IN THE NONPHOSPHORYLATIVE, KETOGENIC OXIDATION
 CC OF GLUCOSE AND OXIDIZES GLUCONATE TO 5-KETOGLUCONATE. DEPENDENT ON
 CC NADP. ALMOST INACTIVE WITH NAD.
 CC -1- CATALYTIC ACTIVITY: D-glucuronate + NAD(P)(+) - 5-dehydro-D-
 CC glucuronate + NAD(P)H.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X80019; CA56322.1; -
 DR HSSP: P47227; 1EDB.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 KW Oxidoreductase; NADP.
 FT NP_BIND 15 39 NADP (BY SIMILARITY).
 FT ACT_SITE 160 160 BY SIMILARITY.
 FT SEQUENCE 236 AA; 27256 MW; 38B03C039C0A07A CRC64;

Query Match 2.9%; Score 9; DB 1; Length 256;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VTGASRGIG 20
 Db 16 VTGASRGIG 24
 |||||

RESULT 12
 ID_FABG_CUPLA STANDARD; PRT; 320 AA.
 AC P28643;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 3-oxoacyl-[acyl-carrier protein] reductase, chloroplast precursor
 DE (EC 1.1.1.100) (3-ketoacyl-acyl carrier protein reductase).
 GN CLKR27.
 OS Cuphea lanceolata.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Myrtales; Lythraceae; Cuphea.
 CC NCB1_TaxID=3930;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92293104; PubMed-1376402;
 RA Klein B., Pawlowski K., Hoerlke-Grandpierre C., Schell J.,
 RA Toepfer R.;
 RT Isolation and characterization of a cDNA from Cuphea lanceolata
 RT encoding a beta-ketoacyl-ACP reductase.*;
 RT Mol. Gen. Genet. 233:122-128(1992).
 RL Mol. Gen. Genet. 233:122-128(1992).
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) - 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
 CC PATHWAY.
 CC -1- SUBUNIT: HOMOTETRAMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST AND NON-PHOTOSYNTHETIC
 CC PLASTIDS.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X64566; CAA5866.1; -
 DR PIR: S19832; S19832.
 DR HSSP: P50162; 1AE1.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 KW Fatty acid biosynthesis; Oxidoreductase; NADP; Chloroplast;

Transit peptide.
 TRANSIT 1 61 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 1 320 3-OXOACYL-(ACYL-CARRIER PROTEIN)
 FT NP_BIND 82 106 REDUCTASE.
 FT ACT_SITE 227 227 NADP (BY SIMILARITY).
 SO SEQUENCE 320 AA; 33103 MW; 06BAF0522B2B8C87 CRC64;

Query Match 2.9%; Score 9; DB 1; Length 320;
 Best Local Similarity 100.0%; Pred. No. 0.19;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 VTGASRGIG 20
 |||||
 DB 82 VTGASRGIG 90

RESULT 13
 Y1V6_YEAST STANDARD; PRT; 263 AA.
 ID Y1V6_YEAST
 AC P40580;
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical oxidoreductase in Y1V1-HYR1 intergenic region
 DE (EC 1.-.-.-).
 DE Y1R036C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA Bartell B.G., Badcock R., Bankier A.T., Bowman S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Lyne G., Moule S., Moule T., Odell C., Pearson D.,
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; Z38061; CAA86196.1; -
 CC PIR; S48498; S48498.
 CC SGD; S0001475; Y1R036C.
 CC InterPro: IPR002198; ADH_short.
 CC Pfam: PF00106; adh_short; 1.
 CC PROSITE: PS00061; ADH_SHORT; 1.
 KW Hypothetical protein; Oxidoreductase.
 FT NP_BIND 6 30 NAD OR NADP (BY SIMILARITY).
 FT ACT_SITE 157 157 BY SIMILARITY.
 SO SEQUENCE 263 AA; 28804 MW; 53DE25425A9C46F CRC64;

Query Match 2.6%; Score 8; DB 1; Length 263;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 TGASRGIG 20
 |||||
 DB 8 TGASRGIG 15

RESULT 14
 DHCA_HUMAN STANDARD; PRT; 276 AA.
 ID DHCA_HUMAN
 AC P6152;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Carbonyl reductase [NADPH] 1 (EC 1.1.1.184) (NADPH-dependent carbonyl
 DE reductase 1) (Prostaglandin-H2 9-reductase) (EC 1.1.1.189)
 DE (Prostaglandin 9-ketoreductase) (15-hydroxyprostaglandin dehydrogenase
 DE [NADPH]) (EC 1.1.1.197).
 DE CBRI OR CBR OR CRN.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Placenta;
 RC MEDLINE=89034082; PubMed=3141401;
 RA Wernuth B., Bohren K.M., Heinemann G., von Wartburg J.-P.,
 RA Gabbay K.H.;
 RT "Human carbonyl reductase. Nucleotide sequence analysis of a cDNA and
 RT amino acid sequence of the encoded protein."
 RL J. Biol. Chem. 263:16185-16188(1988).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Breast;
 RC MEDLINE=90212644; PubMed=2182121;
 RA Forrest G.L., Akman S., Krutzik S., Paxton R.J., Sparkes R.S.,
 RA Doroshov J., Feist R.L., Mohandas T., Bachur N.R.;
 RT "Induction of a human carbonyl reductase gene located on chromosome
 RT 21."
 RL Biochim. Biophys. Acta 1048:149-155(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=92017676; PubMed=1921984;
 RA Forrest G.L., Akman S., Doroshov J., Rivera H., Kaplan W.D.;
 RT "Genomic sequencing and expression of a cloned human carbonyl reductase
 RT gene with daunorubicin reductase activity."
 RL Mol. Pharmacol. 40:502-507(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98414514; PubMed=9740676;
 RA Watanabe K., Sugawara C., Ono A., Fukuzumi Y., Itakura S.,
 RA Yamazaki M., Tashiro H., Osoegawa K., Soeda E., Nomura T.;
 RT "Mapping of a novel human carbonyl reductase, CBR3, and ribosomal
 RT pseudogenes to human chromosome 21q22.2."
 RL Genomics 52:95-100(1998).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Shibuya K., Kudoh J., Minoshima S., Kawasaki K., Nakatoh E.,
 RA Shintani A., Asakawa S., Shimizu N.;
 RT "Genomic sequencing of 1.2-Mb region on human chromosome 21q22.2."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20289799; PubMed=10830953;
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
 RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
 RA Menzel U., Delabar J., Rumpf K., Lehmann R., Patterson D.,
 RA Reichwald A., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordliek G., Hornischer K., Brandt P.,
 RA Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
 RA Ramser J., Beck A., Klages S., Hennig S., Resselmann L., Dagand E.,
 RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
 RA Lehrach H., Reinhardt R., Vaspo M.-L.;
 RT "The DNA sequence of human chromosome 21."
 RL Nature 405:311-319(2000).

```

RN [7]
RP PARTIAL SEQUENCE, AND POST-TRANSLATIONAL MODIFICATIONS.
RX MEDLINE-9313816; PubMed-8421682.
RA Krock M., Ghosh D., Stromberg R., Carlquist M., Joernvall H.;
RT "Carboxyethyllysine in a protein: native carbonyl reductase/NADP(+)-
RT dependent prostaglandin dehydrogenase."
RL Proc. Natl. Acad. Sci. U.S.A. 90:502-506(1993).
CC -1- FUNCTION: CATALYZE THE REDUCTION OF A WIDE VARIETY OF CARBONYL
CC COMPOUNDS INCLUDING THE ANTITUMOR ANTIRACCLINE ANTIOTICS.
CC CAN CONVERT PROSTAGLANDIN E2 TO PROSTAGLANDIN F2-ALPHA.
CC -1- CATALYTIC ACTIVITY: R-CHOH-R + NADP(+) = R-CO-R + NADPH.
CC -1- CATALYTIC ACTIVITY: (5Z,13E)-(15S)-9,11,15-trihydroxyprosta-5,13-
CC -5,13-dienoate + NADPH.
CC -1- CATALYTIC ACTIVITY: (5Z,13E)-(15S)-11,15-dihydroxy-9-oxoprost-
CC -9-oxoprost-13-enoate + NADP(+) = (5Z,13E)-11-alpha-hydroxy-9,15-
CC -dioxoprost-13-enoate + NADPH.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J04056; AAAS2070.1; -
DR EMBL: M62420; AAA17881.1; -
DR EMBL: A3003151; BAA33498.1; -
DR EMBL: AP000688; BAA89424.1; -
DR EMBL: AP001724; BAA95508.1; -
DR PIR: A31912; R0HUCB.
DR PIR: S09013; S09013.
DR HSSP: P14061; 1FDW.
DR MIM: 114830; -
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
DR PRINTS: PR00081; GDHRDH.
DR PROSITE: PS00061; ADH_SHORT; 1.
DR OXidoreductase; NADP; Acetylation.
KW OXidoreductase; NADP; Acetylation.
FT INIT_MET 0
FT MOD_RES 1
FT MOD_RES 238
FT NP_BIND 9
FT ACT_SITE 193
SQ SEQUENCE 276 AA; 30244 MW; 78E83065F5677733 CRC64;

Query Match 2.6%; Score 8; DB 1; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 LDVLYNNA 93
DB 83 LDVLYNNA 90

RESULT 15
FIXR_BRAJA STANDARD; PRT; 278 AA.
AC P05406;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE FIXR protein.
GN FIXR.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Bradyrhizobium.

```

```

OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88040468; PubMed-3313281;
RA Thoeny B., Fischer H.-W., Anthamatten D., Bruderer T., Hennecke H.;
RT "The symbiotic nitrogen fixation regulatory operon (fixRnifA) of
RT Bradyrhizobium japonicum is expressed aerobically and is subject to a
RT novel, nifA-independent type of activation."
RL Nucleic Acids Res. 15:8479-8499(1987).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X06167; CAA29530.1; -
DR PIR: S01065; S01065.
DR HSSP: P08074; 1CYD.
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; 1.
KW Nitrogen fixation; Oxidoreductase.
FT NP_BIND 40
FT ACT_SITE 189
SQ SEQUENCE 278 AA; 29715 MW; 69FDC702072A20A0 CRC64;

Query Match 2.6%; Score 8; DB 1; Length 278;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TGASRGIG 20
DB 42 TGASRGIG 49

Search completed: July 31, 2002, 15:17:40
Job time: 223 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

```
Run on:      July 31, 2002, 15:05:01 ; Search time 30.31 Seconds
              (without alignments)
              1147.017 Million cell updates/sec
```

Title: US-10-006-163-1
Page: 1633

Sequence: 1 MAAPMNGVCVVTGASRGIG.....YLPSELRVPKWIIALYTSKF 313

Scoring table: BLOSUM62

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : A_Geneseq_032802:*

1:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1980.DAT.*
2:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1981.DAT.*
3:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1982.DAT.*
4:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1983.DAT.*
5:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1984.DAT.*
6:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1985.DAT.*
7:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1986.DAT.*
8:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1987.DAT.*
9:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1988.DAT.*
10:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1989.DAT.*
11:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1990.DAT.*
12:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1991.DAT.*
13:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1992.DAT.*
14:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1993.DAT.*
15:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1994.DAT.*
16:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1995.DAT.*
17:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1996.DAT.*
18:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1997.DAT.*
19:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1998.DAT.*
20:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA1999.DAT.*
21:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA2000.DAT.*
22:	/SIDSI/gcgdata/hold-genseeq/genseeq-emb1/AA2001.DAT.*

SUMMARIES

Query No.	Score	Match	Length	DB	ID	Description
1	1623	100.0	313	20	AAV27004	Human short-chain
2	1619	99.8	313	21	AA842640	Human ORF2404
3	1619	99.8	313	22	AA881051	Human protein HP01
4	746	46.0	162	22	AA873423	Human gene 22-enco
5	361	22.9	127	22	AA873479	Human gene 22-enco
6	371	22.2	72	22	AA873481	Human gene 22-enco
7	344	21.2	303	22	AA895602	Human protein sequ
8	273.45	16.9	258	21	AA856815	Bacillus D-arabini
9	267	16.5	122	22	AA873483	Human gene 22-enco
10	266	15.4	68	22	AA873485	Human gene 22-enco
11	253	15.6	248	21	AA154422	Amino acid sequenc

12	227.5	13.7	257	22	AB559275	Drosophila melanog
13	228.5	13.7	256	17	AA062111	Glucuronate:NADP+-5-
14	228.5	13.7	315	17	AA089323	Rape leaf beta-Xet
15	228.5	13.7	315	17	AA089322	Rape seed beta-Xet
16	214.5	13.2	246	21	AA015707	Staphylococcus aur
17	214.5	13.2	246	22	AA033965	Staphylococcus aur
18	214.5	13.2	246	22	AA036530	Staphylococcus aur
19	214.5	13.2	246	22	AA037210	Staphylococcus aur
20	214.5	13.2	246	22	AA037507	Staphylococcus aur
21	212.5	13.1	246	22	AA021957	S. aureus NADPH-de
22	211.5	13.0	263	21	AA064729	Arabidopsis thalia
23	211.5	13.0	263	21	AA051431	Arabidopsis thalia
24	211.5	13.0	273	21	AA051430	Arabidopsis thalia
25	211.5	13.0	274	21	AA064728	Arabidopsis thalia
26	210	12.9	240	21	AA0605872	Arabidopsis thalia
27	210	12.9	242	21	AA0605871	Arabidopsis thalia
28	207.5	12.8	251	22	AA034193	Staphylococcus aur
29	207.5	12.8	272	22	AA037095	Staphylococcus aur
30	207	12.8	412	22	AB061661	Drosophila melanog
31	206.5	12.7	263	21	AA0605381	Arabidopsis thalia
32	206.5	12.7	279	21	AA0605384	Arabidopsis thalia
33	203.5	12.3	279	21	AA0605383	Arabidopsis thalia
34	203.5	12.5	290	21	AA0605462	Arabidopsis thalia
35	203.5	12.5	1639	22	AB055281	Drosophila melanog
36	202.5	12.5	438	22	AA081258	Human AFP protein
37	202.5	12.5	418	22	AA081260	Human AFP protein
38	202.5	12.5	418	22	AA081367	Amino acid sequenc
39	199.5	12.3	267	21	AA063505	Arabidopsis thalia
40	199.5	12.3	285	21	AA0612862	Arabidopsis thalia
41	199.5	12.3	307	21	AA0612861	Arabidopsis thalia
42	199.5	12.3	308	21	AA063504	Arabidopsis thalia
43	197.5	12.2	285	21	AA060333	Arabidopsis thalia
44	197.5	12.2	307	21	AA060332	Arabidopsis thalia
45	197	12.1	357	22	AA018345	Human endocrine po

ALIGNMENTS

RESULT	1
AAZ27004	
ID	AAZ27004 standard; Protein: 313 AA.
XX	
AC	AAZ27004;
XX	
DT	20-SEP-1999 (first entry)
XX	
DE	Human short-chain dehydrogenase (HSCD) enzyme.
XX	
KE	Short-chain dehydrogenase; HSCD; enzyme; pyruvate; coenzyme A; human;
KW	mitochondria; immune disorder; cancer; leukemia; adenocarcinoma;
KW	lymphoma; breast; lung; testis; prostate; brain; Addison's disease;
KW	acquired immune deficiency syndrome; asthma; anemia; Crohn's disease;
KW	Graves disease; AIDS; gene therapy.
XX	
OS	Homo sapiens.
XX	
PN	US5928923-A.
XX	
PD	27-JUL-1999.
XX	
PF	05-FEB-1998; 98US-0019216.
XX	
PR	05-FEB-1998; 98US-0019216.
XX	
PA	(INCY-) INCYTE PHARM INC.
XX	
PI	Corley NC, Lai P;
XX	
DR	WPI: 1999-429503/36.
DR	N-PSDB: AAX86772.
XX	
PT	Nucleic acids encoding human short chain dehydrogenase enzymes

PT useful for the diagnosis, prevention and treatment of disorders
 PT associated with abnormal expression of the enzyme, such as immune
 PT disorders and cancer

PS Claim 1: Fig 1A-D; 27pp; English.

XX This represents a human short-chain dehydrogenase (HSD) enzyme. The
 CC dehydrogenase enzymes catalyse an irreversible reaction between pyruvate
 CC and coenzyme A, to form CO₂ and the intermediate CoA, in mitochondria.
 CC Host cells containing vectors comprising the HSD nucleic acid may be
 CC used to produce the HSD enzyme, according to standard recombinant DNA
 CC methodology. The enzyme may then be used as an antigen in the production
 CC of antibodies or in assays to identify antagonists of HSD activity.
 CC These antagonists may then be used to treat disorders associated with
 CC inappropriate expression, or over activity of HSD such as immune
 CC disorders and cancers (the antagonist interferes with the reaction
 CC between pyruvate and coenzyme A). For example, the antagonists may be
 CC used to treat leukemia, lymphomas, adenocarcinomas and cancers of the
 CC breast, lung, testis, prostate and brain, Addison's disease, acquired
 CC immune deficiency syndrome (AIDS), asthma, anemia, Crohn's disease and
 CC Graves disease. The nucleic acids and antisense sequences can be used
 CC in gene therapy.

XX Sequence 313 AA:

Query Match 100.0%; Score 1623; DB 20; Length 313;
 Best Local Similarity 100.0%; Pred. No. 5.1e-164;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPNMGVCVVTGASRGIGRGIALQLCAGATVYITGRHDLTLRVVAQEAOSLGQCYP 60
 DB 1 maapnmgvcvvtgatsrgigrgialqlckagatvyltgrhdtltlvvgeaqs199qcvp 60
 QY 61 VVCDSSQSESEVRLTEFQVDRQOGRDLVYNNAVAGVOTILTRNKAKEMETPASMWDIN 120
 DB 61 vvcdssqsesevrltefqvdrqogrldvlynnavagvotlitrnkakemetpsamwddin 120
 QY 121 NVGLGHHFCSVYGARLWVPAOGGLIVYSSPSGQYMFNVPYGVGKACDKLADCAHE 180
 DB 121 nvglgghfcsvygarlwpagggllvlysspslqymfnvpygvkaccdklaadcahe 180
 QY 181 LRRHGVSCVSLMPGIVOTELKEHNAKEEVLQDPVLKQFSAFSAETTELSGKCVVLA 240
 DB 181 lrhgvscvslmpgivotellkehnakeevlqdpvlkqfsafsaettelesgkcvvlla 240
 QY 241 TDPNLLISGKVLPCSDLARRIGLFDVGRPVQDYLSSVLSHVSGLGWLASLPSELR 300
 DB 241 tdpnllisgkvlpcsdliarrylrdvgrpvqdyllssvlsahvgslgwlasylpsflr 300
 QY 301 VPKWIIALYTSKF 313
 DB 301 vpkwiialytskf 313

RESULT 2

AAB42640
 ID AAB42640 standard; Protein: 313 AA.

XX AAB42640;

DT 08-FEB-2001 (first entry)

XX Human ORFX ORF2404 polypeptide sequence SEQ ID NO:4808.

XX Human: open reading frame; ORFX; detection: cytosolic; hepatotropic;
 KW vulnary; antipsoatic; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;
 KW antiviral; antibacterial; antifungal; antineumatic; antitumor;
 KW antinaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX Homo sapiens.

XX MO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shinketsu RA, Leach M;

XX WPI: 2000-602362/57.

XX N-PSDB: AAC76849.

XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 3990; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytosolic; hepatotropic; vulnary;
 CC antipsoatic; antiparkinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antifungal; antineumatic;
 CC antitumor; and antinaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 313 AA:

Query Match 99.8%; Score 1619; DB 21; Length 313;
 Best Local Similarity 99.7%; Pred. No. 1.3e-163;
 Matches 312; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPNMGVCVVTGASRGIGRGIALQLCAGATVYITGRHDLTLRVVAQEAOSLGQCYP 60
 DB 1 maapnmgvcvvtgatsrgigrgialqlckagatvyltgrhdtltlvvgeaqs199qcvp 60
 QY 61 VVCDSSQSESEVRLTEFQVDRQOGRDLVYNNAVAGVOTILTRNKAKEMETPASMWDIN 120
 DB 61 vvcdssqsesevrltefqvdrqogrldvlynnavagvotlitrnkakemetpsamwddin 120
 QY 121 NVGLGHHFCSVYGARLWVPAOGGLIVYSSPSGQYMFNVPYGVGKACDKLADCAHE 180
 DB 121 nvglgghfcsvygarlwpagggllvlysspslqymfnvpygvkaccdklaadcahe 180

```

QY 181 LRRHGVCSVSLMPGIVOTELLKEHMAKEEVLQDPVLKQFKSAFSSAETTELSGKCVALA 240
DB 181 Lrrhgvacsvslmpgivgtellkehmakeevlqdpvlkqfkfsaetfelsgcvuala 240
QY 241 TDPNLTLSGKVLPSCDLARRYGLRDVGRPVQDYLSSVLSHNSGIMGLASTLPSTLR 300
DB 241 tdpnltslsgkvlpscdlarrylrdvgrpvqdyllssvlsnsglglwlaasyllpstlr 300
QY 301 VPKWIALYTSKF 313
DB 301 vpkwialytskf 313

RESULT 3
AAB81051
ID AAB81051 standard; protein: 313 AA.
AC AAB81051;
DE 20-JUN-2001 (first entry)
XX Human protein HP01017 amino acid sequence.
XX Homo sapiens.
XX JP2001037482-A.
XX 13-FEB-2001.
XX 28-JUL-1999; 99JP-0214315.
XX 28-JUL-1999; 99JP-0214315.
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX WPI; 2001-285599/30.
XX N-PSDB; AAF77480.
XX Novel protein useful for the detection of a receptor and a ligand, and
XX for screening low molecular weight drugs.
XX Claim 1; Fig 5; 35pp; Japanese.
XX This invention relates to purified human proteins AAB81047 - AAB81056
XX which are encoded by cDNA sequences AAF77476 - AAF77485. The invention
XX includes an expression vector which can translate DNA encoding the
XX protein or express it in a cell. Also included is a cell transformed by
XX the vector, and an antibody specific for the protein. The protein can be
XX used for the detection of receptors and ligands and in the screening for
XX new low molecular weight drugs. The present sequence represents the
XX protein product of the human cDNA clone HP00758.
XX Sequence 313 AA;

Query Match 99.8%; Score 1619; DB 22; Length 313;
Best Local Similarity 99.7%; Pred. No. 1.3e-163;
Matches 312; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 181 LRRHGVCSVSLMPGIVOTELLKEHMAKEEVLQDPVLKQFKSAFSSAETTELSGKCVALA 240
DB 181 Lrrhgvacsvslmpgivgtellkehmakeevlqdpvlkqfkfsaetfelsgcvuala 240
QY 241 TDPNLTLSGKVLPSCDLARRYGLRDVGRPVQDYLSSVLSHNSGIMGLASTLPSTLR 300
DB 241 tdpnltslsgkvlpscdlarrylrdvgrpvqdyllssvlsnsglglwlaasyllpstlr 300
QY 301 VPKWIALYTSKF 313
DB 301 vpkwialytskf 313

RESULT 4
AAG73423
ID AAG73423 standard; Protein: 162 AA.
AC AAG73423;
DE 10-AUG-2001 (first entry)
XX Human gene 22-encoded secreted protein HHFC267, SRQ ID NO:195.
XX Human; secreted protein; proliferative disorder; cancer; chromosome 14;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
XX cardiovascular disorder; angiogenic disorder; kidney disorder;
XX gastrointestinal disorder; pregnancy-related disorder; tumour;
XX endocrine disorder; infection; wound healing; vulnery;
XX cell culture; chemotaxis; food additive;
XX binding partner identification.
XX Homo sapiens.
XX WO200134628-A1.
XX 17-MAY-2001.
XX 08-NOV-2000; 2000WO-US30653.
XX 12-NOV-1999; 99US-0164735.
XX 27-JUL-2000; 2000US-0221193.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Komatsu S, GA, Birse CE, NI J, Moore PA;
XX WPI; 2001-329066/34.
XX N-PSDB; AAH32600.
XX Nucleic acids encoding 35 human secreted polypeptides, useful for
XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's
XX disease and diabetic retinopathy.
XX Claim 11; Page 549; 604pp; English.
XX AAH2522-AAH3267 represent cDNAs corresponding to 35 human secreted
XX protein genes; and AAG73448 represent the proteins they encode.
XX AAG73449-AAG73519 represent human secreted protein fragments. The genes
XX and their corresponding secreted proteins are useful for preventing,
XX treating or ameliorating medical conditions, e.g., by protein or gene
XX therapy. Pathological conditions can be diagnosed by determining the
XX amount of the new protein in a sample or by determining the presence of
XX mutations in the new genes. Specific uses are described for each of the
XX 52 genes, based on the tissues in which they are most highly expressed,
XX and include developing products for the diagnosis or treatment of
XX proliferative disorders, cancer, tumours, foetal and developmental
XX abnormalities, haematopoietic disorders, diseases of the immune system,
XX AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,

```


CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiotensin, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein of the invention.
 XX
 SQ Sequence 162 AA;

Query Match 46.0%; Score 746; DB 22; Length 162;
 Best Local Similarity 100.0%; Pred. No. 4.6e-71;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 CDKLAADCAHELRHGVCSVSLMPGIVOTELKEHMAKEEVLQDPVLKOPKSAFSAET 229
 Db 19 cdklaadcahelrrhgvcsvslmpgivotellkemaeevlqdpvlkqfksaissaet 78
 QY 230 ELSGCYVALNDPRLISLGSKVLPSCDLARRYGLRDVGRPYDYLSSVLSHVSGLG 289
 Db 79 elsgcyvalndprrlislsgskvlpscdlarryglrdvgrpydyllssvlsahvs9lg 138
 QY 290 WLAATPSFLRVPKVIILYTSKF 313
 Db 139 wlaatpsflrvpkviallytskf 162

RESULT 5
 AAG73479
 ID AAG73479 standard; Protein; 127 AA.
 AC AAG73479;

DT 10-AUG-2001 (first entry)
 DE Human gene 22-encoded secreted protein fragment; SEQ ID NO:254.
 XX

KM Human; secreted protein; proliferative disorder; cancer; chromosome 14;
 KM foetal abnormality; developmental abnormality; haematopoietic disorder;
 KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KM inflammation; allergy; neurological disorder; Alzheimer's disease;
 KM Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KM skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KM cardiovascular disorder; angiotensin; kidney disorder;
 KM gastrointestinal disorder; pregnancy-related disorder; tumour;
 KM endocrine disorder; infection; wound healing; vulnerability;
 KM cell culture; chemotaxis; food additive;
 KM binding partner identification.

OS Homo sapiens.
 XX
 XX
 PN WO200134628-A1.
 PD 17-MAY-2001.
 XX

PF 08-NOV-2000; 2000MO-US30653.
 XX
 XX 12-NOV-1999; 99US-0164735.
 PR 27-JUL-2000; 2000US-0221193.
 XX

XX (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Komatsoulis GA, Birse CE, NI J, Moore PA;
 PI

XX
 DR WPI: 2001-329066/34.
 XX

PT Nucleic acids encoding 35 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 PS Disclosure; Page 47; 604pp; English.

CC AAH3252-AAH3262 represent cDNAs corresponding to 35 human secreted
 CC protein genes, and AAG73448-AAG73449 represent the proteins they encode.
 CC AAG73449-AAG73519 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 52 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., Rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiotensin, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.
 XX
 SQ Sequence 127 AA;

Query Match 22.9%; Score 371; DB 22; Length 127;
 Best Local Similarity 63.1%; Pred. No. 2.8e-31;
 Matches 82; Conservative 6; Mismatches 12; Indels 30; Gaps 4;

QY 193 PGIVOTELKEHMAKEEVLQDPVLKOPKSAFSS-AETTELSGKV-----VALATDP 243
 Db 19 pgiivotellkemaeevlqdpvlkopsafss-aettellsgkv-----valatdp 57
 QY 244 NILSLGKVLPSCDLARRYGLRDVGRPYDYLSSVLSHVSGLGWLASTLPSTLRPK 303
 Db 58 nilslgkvlpscdlarryglrdvgrpydyllssvlsahvs9lgwlaatpsflrvpk 117
 QY 304 WITIALYTSKF 313
 Db 118 witalnyskf 127

RESULT 6
 AAG73481
 ID AAG73481 standard; Protein; 72 AA.
 AC AAG73481;

DT 10-AUG-2001 (first entry)

DE Human gene 22-encoded secreted protein fragment, SEQ ID NO:256.

KM Human; secreted protein; proliferative disorder; cancer; chromosome 14;
 KM foetal abnormality; developmental abnormality; haematopoietic disorder;
 KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KM inflammation; allergy; neurological disorder; Alzheimer's disease;
 XX
 XX
 XX

KM Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KM skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KM cardiovascular disorder; angiodysplasia; kidney disorder;
 KM gastrointestinal disorder; pregnancy-related disorder; tumour;
 KM endocrine disorder; infection; wound healing; vulnerability;
 KM cell culture; chemotaxis; food additive;
 KM binding partner identification.
 KM
 OS Homo sapiens.
 PN MO200134628-A1.
 XX
 PD 17-MAY-2001.
 XX
 PF 08-NOV-2000; 2000MO-US30653.
 XX
 PR 12-NOV-1999; 99US-0164735.
 PR 27-JUL-2000; 2000US-0221193.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ruben SM, Komatsoulis GA, Birse CE, Ni J, Moore PA;
 DR WPI; 2001-329066/34.
 XX
 PT Nucleic acids encoding 35 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 XX
 PS Disclosure; Page 47; 604pp; English.
 XX
 CC AAH32522-AAH32627 represent cDNAs corresponding to 35 human secreted
 CC protein genes, and AAH32346-AAH32448 represent the proteins they encode;
 CC AAH32346-AAH32351 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 32 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.
 XX
 SQ Sequence 72 AA:
 Query Match 22.2%; Score 361; DB 22; Length 72;
 Best Local Similarity 97.2%; Pred. No. 1.3e-30;
 Matches 70; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 61 pkwialnatski 72
 ||||| |||||
 RESULT 7
 ID AAB95802
 ID AAB95802 standard; Protein: 303 AA.
 AC AAB95802;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:18783.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KW Homo sapiens.
 OS
 XX
 PM EP1074617-A2.
 PD 07-FEB-2001.
 PF 28-JUL-2000; 2000EP-0116126.
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-018776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 18783; 2537bp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 303 AA:
 Query Match 21.2%; Score 344; DB 22; Length 303;
 Best Local Similarity 34.4%; Pred. No. 8.4e-28;
 Matches 120; Conservative 35; Mismatches 92; Indels 102; Gaps 19;

Query Match	16.9%	Score 273.5;	DB 21;	Length 258;
Best Local Similarity	34.8%	Pred. No. 2.1e-20;		
Matches	72;	Conservative 33;	Mismatches 87;	Indels 15;
			Gaps	5

RESULT	9	
AAAG73483	standard; Protein; 122 AA.	
AAAG73483		
AAAG73483		
10-AUG-2001	(first entry)	
Human gene 22-encoded secreted protein fragment, SEQ ID NO:258.		
Human; secreted protein; proliferative disorder; cancer; chromosome 14;		
fetal abnormality; developmental abnormality; haematopoietic disorder;		
immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;		
inflammation; allergy; neurological disorder; Alzheimer's disease;		
Parkinson's disease; cognitive disorder; schizophrenia; asthma;		
skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;		
cardiovascular disorder; angio-genic disorder; kidney disorder;		
gastrointestinal disorder; pregnancy-related disorder; tumour;		
endocrine disorder; infection; wound healing; vulnary;		
cell culture; chemotaxis; food additive;		
binding partner identification.		
Homo sapiens.		
MO200134628-A1.		
17-MAY-2001.		
08-NOV-2000; 2000MO-US30653.		
12-NOV-1999; 99US-0164735.		
27-JUL-2000; 2000US-0221193.		
(HUMA-) HUMAN GENOME SCI INC.		
Ruben SM, Komatsu-GA, Blise CE, NI J, Moore PA;		
WPI; 2001-329066/34.		
Nucleic acids encoding 35 human secreted polypeptides, useful for		
preventing, diagnosing and/or treating e.g. cancers, Parkinson's		
disease and diabetic retinopathy -		
Disclosure: Page 47; 604pp; English.		
AAH32522-AAH32627 represent cDNAs corresponding to 35 human secreted		
protein genes, and AAAG73446-AAAG73448 represent the proteins they encode		
AAAG73449-AAAG73519 represent human secreted protein fragments. The genes		
and their corresponding secreted proteins are useful for preventing,		
treating or ameliorating medical conditions, e.g., by protein or gene		
therapy. Pathological conditions can be diagnosed by determining the		
amount of the new protein in a sample or by determining the presence of		
mutations in the new genes. Specific uses are described for each of the		
52 genes, based on the tissues in which they are most highly expressed,		
and include developing products for the diagnosis or treatment of		
proliferative disorders, cancer, tumours, foetal and developmental		
abnormalities, haematopoietic disorders, rheumatoid arthritis, diabetes,		
asthma, Parkinson's disease, Alzheimer's disease, schizophrenia, skin		
disorders, cardiovascular disorders, angio-genic disorders, kidney disorders,		
gastrointestinal disorders, pregnancy-related disorders, tumour;		
endocrine disorders, infection, wound healing; vulnary;		
cell culture; chemotaxis; food additive;		
binding partner identification.		
Homo sapiens.		
MO200134628-A1.		
17-MAY-2001.		
08-NOV-2000; 2000MO-US30653.		
12-NOV-1999; 99US-0164735.		
27-JUL-2000; 2000US-0221193.		
(HUMA-) HUMAN GENOME SCI INC.		
Ruben SM, Komatsu-GA, Blise CE, NI J, Moore PA;		
WPI; 2001-329066/34.		
Nucleic acids encoding 35 human secreted polypeptides, useful for		
preventing, diagnosing and/or treating e.g. cancers, Parkinson's		
disease and diabetic retinopathy -		
Disclosure: Page 47; 604pp; English.		

CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiotensin disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein fragment referred to in the disclosure of the invention.
XX
SO Sequence 122 AA:

Query Match 16.5%; Score 267; DB 22; Length 122;
Best Local Similarity 54.9%; Pred. No. 3.1e-20;
Matches 62; Conservative 8; Mismatches 23; Indels 20; Gaps 5;

OY 170 CDKLADCAHELRHGVCSVLMFGIVOTELKEMAKEVYLDVPLKQF-KSAFSSAET 228
DB 12 cdklaadcahelelrhgvcsvslmpgltvgtellkemaakevldqpvtkyqkgrakeeen 71

OY 229 TELSGKCVVALDPNLTLSGKVLPS-----CDLAR-RVGLNDVGRPYOD 274
DB 72 rgyv-gldc-----lnkrlryshspgagycgrcraengthllsfnpvre 112

RESULT 10
AAG73485
ID AAG73485 standard; Protein: 68 AA.

AC AAG73485;

DT 10-AUG-2001 (first entry)

DE Human gene 22-encoded secreted protein fragment, SEQ ID NO:260.

XX Human; secreted protein; proliferative disorder; cancer; chromosome 14;
XX foetal abnormality; developmental abnormality; hematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
XX cardiovascular disorder; angiotensin disorder; kidney disorder;
XX gastrointestinal disorder; pregnancy-related disorder; tumour;
XX endocrine disorder; infection; wound healing; vulneryary;
XX cell culture; chemotaxis; food additive;
XX binding partner identification.

OS Homo sapiens.

PN NC200134628-A1.

PD 17-MAY-2001.

PF 08-NOV-2000; 2000MO-US30653.

PR 12-NOV-1999; 99US-0164735.

XX 27-JUL-2000; 2000US-0221193.

PA (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Komatsoulis GA, Birse CE, NI J, Moore PA;
XX WPI, 2001-3329066/34.
XX

PT Nucleic acids encoding 35 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
XX
PS Disclosure; Page 47; 604pp; English.

XX AAH32522-AAH3267 represent cDNAs corresponding to 35 human secreted
XX protein genes, and AAG73346-AAG73448 represent the proteins they encode.
XX AAG73449-AAG73519 represent human secreted protein fragments. The genes
XX and their corresponding secreted proteins are useful for preventing,
XX treating or ameliorating medical conditions, e.g., by protein or gene
XX therapy. Pathological conditions can be diagnosed by determining the
XX amount of the new protein in a sample or by determining the presence of
XX mutations in the new genes. Specific uses are described for each of the
XX 52 genes, based on the tissues in which they are most highly expressed,
XX and include developing products for the diagnosis or treatment of
XX proliferative disorders, cancer, tumours, foetal and developmental
XX abnormalities, hematopoietic disorders, diseases of the immune system,
XX AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
XX allergies, neurological disorders (e.g., Alzheimer's disease,
XX Parkinson's disease), cognitive disorders, schizophrenia, asthma,
XX skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
XX cardiovascular disorders, angiotensin disorders, kidney disorders,
XX gastrointestinal disorders, pregnancy-related disorders, endocrine
XX disorders, and infections. The proteins can also be used to aid wound
XX healing and epithelial cell proliferation, to prevent skin aging due to
XX sunburn, to maintain organs before transplantation, for supporting cell
XX culture of primary tissues, to regenerate tissues, to identify their
XX cognate ligands or binding partners, and in chemotaxis, and can be used
XX as a food additive or preservative to modify storage properties.
XX Antibodies specific for a protein of the invention can be used in
XX alleviating symptoms associated with the disorders mentioned above, and
XX in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
XX immunosorbent assay (ELISA). The present sequence represents a human
XX secreted protein fragment referred to in the disclosure of the invention.
XX
SO Sequence 68 AA:

Query Match 16.4%; Score 266; DB 22; Length 68;
Best Local Similarity 78.3%; Pred. No. 1.6e-20;
Matches 54; Conservative 3; Mismatches 10; Indels 2; Gaps 2;

OY 170 CDKLADCAHELRHGVCSVLMFGIVOTELKEMAKEVYLDVPLKQF-KSAFSSAET 228
DB 1 cdklaadcahelelrhgvcsvslmpgltvgtellkemaakevldqpvtkyqkgrakeeen 60

OY 229 TELSGKCVV 237
DB 61 rgyv-gldc 68

RESULT 11

ID AAY54422 standard; Protein: 248 AA.

AC AAY54422;

DT 06-APR-2000 (first entry)

DE Amino acid sequence of a beta-ketoacyl-ACP reductase protein.

XX Beta-ketoacyl-ACP reductase; Type II fatty acid synthetase;
XX steroselectivity; 4-chloroacetate acid ester;
XX (S)-4-halo-3-hydroxybutyric acid ester; 4-halo-acetoacetic acid ester;
XX beta-ketoacyl-acyl carrier protein reductase; acetoacetyl-CoA reductase;
XX polybeta-hydroxy fatty acid biosynthesis; optically active;
XX 4-naol-3-hydroxybutyric acid ester.
XX
XX Bacillus subtilis.
XX
XX EP955375-A2.

```

PD 10-NOV-1999.
PF 10-MAY-1999; 99EP-0109403.
PR 08-MAY-1998; 98JP-0126507.
PR 21-OCT-1998; 98JP-0300178.
PR 05-APR-1999; 99JP-0098205.
XX
PA (DAIL ) DAICEL CHEM IND LTD.
XX
XX Yamamoto H;
PI
PI
DR MPI; 2000-118183/11.
DR N-PSDB; AA245749.
XX
XX Producing optically pure (S)-4-halo-3-hydroxybutyric acid ester -
PS Claim 6; Page 19-20; 34pp; English.
XX
XX
XX The present sequence represents a beta-ketoacyl-ACP reductase protein
XX of Bacillus subtilis. The beta-ketoacyl-ACP reductase enzyme constitutes
XX a type II fatty acid synthetase. The enzyme has an extremely high
XX reducing activity and stereoselectivity towards 4-chloroacetoacetic
XX acid ester. The specification describes a method for producing a
XX (S)-4-halo-3-hydroxybutyric acid ester. The method comprises
XX asymmetrically reducing 4-halo-acetoacetic acid ester or its
XX derivative with beta-ketoacyl-acyl carrier protein reductase
XX constituting type II fatty acid synthase, or acetoacetyl-CoA
XX reductase constituting the polybeta-hydroxy fatty acid biosynthesis
XX system. The novel method is used to produce optically active
XX 4-hal-3-hydroxybutyric acid ester, with a high purity.
XX
XX Sequence 248 AA;
SQ

```

OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001MO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PMD, Myers EW;
XX
XX WPI: 2001-656860/75.
DR N-PSDB; ABL03378.
XX
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
PS Disclosure; SEQ ID NO 4617; 21pp + Sequence Listing; English.
XX
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins
CC sequences (AB101840-AB16175).
CC (AB157737-AB172072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 257 AA;

[illegible]

XX Gluconate; NADP: oxidoreductase; Gluconobacter oxydans;
KM ascorbic acid; tartaric acid; ds.
XX Gluconobacter oxydans.
XX EP726320-A2.
XX 14-AUG-1996.
XX 07-FEB-1996; 96EP-0101776.
XX 07-FEB-1995; 95DE-4003946.
XX (RHEI-) RHEIN BIOTECH GES NEUE BIOTECHNOLOGISCHE.
XX Bringer-Meyer S, Hollenberg CP, Klasen R, Salm H;
XX MPI: 1996-364397/37.
XX N-PSDB: AAT36145.
XX 5-keto:gluconate prodn. by increasing expression of gluconate:
PT NADP+-5-oxido:reductase gene - esp. by increasing copy no. in
PT Gluconobacter, used as intermediate for ascorbic and tartaric acids
XX
PS Claim 8; Page 10-12; 15pp; German.
XX The gluconobacter oxydans gene may be used to transform cells,
CC to produce higher levels of gluconate:NADP+-5-oxido:reductase.
CC Tartaric acid can now be produced without the difficult
CC purification involved in prepn. from tartar.
XX
SQ Sequence 256 AA:

Query Match 13.7%; Score 222.5; DB 17; Length 256;
Best Local Similarity 31.1%; Pred. No. 5.6e-15;
Matches 66; Conservative 32; Mismatches 93; Indels 21; Gaps 5;

QY 5 MNGQVCVTGASRGIGRGIALOLCRAGATVYITGRHLDTLRVAQEOASLGQCVPVCD 64
Db 9 lsgaralvtgsargyltlakglarygeevlmgnaesldasqfseaglkastavtd 68
QY 65 SSQSESVRTLEQVDREOQGRDLVNNAYAGVOTILNTRNKAFTWTPASMDINNVL 124
Db 69 vldqgavldgvaalerd-mgpidillm--agiq-----trpleefsrkdwddlmstnv 120
QY 125 RGHYFCSTYGARLWVPAGOGILIVISS-----PGSLQYMFNVPRYGVGAKCDKLADG 177
Db 121 navftvgavahmiprgyrgkivnlcsyqselarpqi-----apyratkgaavknltkqm 174
QY 178 AHELRRHGVCVSLMPGIVOTELKEHMAKEE 209
Db 175 atdwgrhlgqlnglapyfatemterlvadee 206

RESULT 14
AAR89323
ID AAR89323 standard; Protein; 315 AA.
XX
AC AAR89323;
XX
DT 13-APR-1996 (first entry)
XX
DE Rape leaf beta-ketoacyl-ACP-ketoreductase.
XX
KM Rape; leaf; beta-ketoacyl-ACP-ketoreductase; beta-ketoreductase;
KM plasmid pURL6.2; cDNA library; Escherichia coli; vector;
KM plasmid; stroma; transit peptide; cassette; antisense; oilseed;
KM transgenic plant; crop improvement; lipid; metabolic engineering;
KM polymer; rapeseed oil.
XX
OS Brassica napus.

XX Key Location/Qualifiers
FT Peptide 1..55 "transit peptide"
XX
XX W09602652-A2.
XX 01-FEB-1996.
XX 17-JUL-1995; 95WO-GB01678.
XX 20-JUL-1994; 94GB-0014622.
XX (ZENEC) ZENECA LTD.
XX Chase D, Elborough K, Fentem PA, Slabas AR, White A;
XX MPI: 1996-105914/11.
XX N-PSDB: AAG99305.
XX
XX New isolated rape beta-ketoreductase DNA - used to develop plants
PT with lower or higher oil contents or with altered oil compsn.
XX
PS Claim 2; Page 16; 29pp; English.
XX
CC The sequence corresponds to a rape leaf beta-ketoreductase encoded
CC by a cDNA insert in plasmid pURL6.2 in Escherichia coli XL1-Blue.
CC A plastid stroma targeting transit peptide is present. DNA
CC encoding the protein may be inserted in a vector or expression
CC cassette in sense or antisense orientation for expression in an
CC oilseed plant, e.g. for production of transgenic rape plants with
CC low or modified oil content, diversion of metabolism to alternative
CC storage compounds, e.g. starch, protein or engineered polymers, or
CC production of plants with enhanced oil content. The DNA may also
CC be used as a probe to obtain similar genes from other plants. The
CC transit peptide may be used to direct other proteins to leaf
CC plastids.
XX
SQ Sequence 315 AA:

Query Match 13.7%; Score 222.5; DB 17; Length 315;
Best Local Similarity 25.7%; Pred. No. 7.9e-15;
Matches 75; Conservative 50; Mismatches 108; Indels 59; Gaps 9;

QY 2 AAF-MNGQVCVTGASRGIGRGIALOLCRAGATVYITGRHLDTLRVAQEOASLGQCVCV 59
Db 66 avpkvespvyvvvtgargykalalslgkagckvlnyarsakeaeveksqleaygqai 125
QY 60 PIVCDSSQSESVRTLEQVDREOQGRDLVNNAYAGVOTILNTRNKAFTWTPASMDINNVL 119
Db 126 tfggdvskdeaveamkta-ldawgtldvvn--agj-----trdclllrmkksqwdv 177
QY 120 NNWGLGHHFCSYTGARLWVPAGOGILIVISS--PGSLQYMFNVPRYGVGAKCDKLADG 178
Db 178 ldnltgyvltcgaacklmkkkrkgrlfnlasvvgllgnlgqanysaaagvlgfstka 237
QY 179 HELRRHGVCVSLMPGIVOTEL---LKEHMAKEEVLDPLVKOFKAFSAETTELSGKC 235
Db 238 regaarnlnvnyvcpgfiasdmaklgedmek----- 269
QY 236 VVALATDPNILLSGRVLPSCDLARRYGRLDVGPRVQDYLSSVLSVAVSG 287
Db 270 -----klgltpigrigypqevag--lveflaipaasyltg 304

RESULT 15
AAR89322
ID AAR89322 standard; Protein; 315 AA.
XX
AC AAR89322;
XX
DT 13-APR-1996 (first entry)

XX Rape seed beta-ketoacyl-ACP-ketoreductase.

XX Rape; seed: beta-ketoacyl-ACP-ketoreductase; beta-ketoreductase;
KM plasmid: pJRS10.1; CDNA library: embryo; Escherichia coli; vector;
KM plasmid: strona; transit peptide; cassette; antisense; oilseed;
KM transgenic plant; crop improvement; lipid; metabolic engineering;
KM polymer; rapeseed oil.

OS Brassica napus.

Key	Location/Qualifiers
FT Peptide	1..56
FT	/note- "Transit peptide"

XX WO9602652-A2.

XX 01-FEB-1996.

XX 17-JUL-1995; 95WO-GB01678.

XX 20-JUL-1994; 94GB-0014622.

XX (ZENECA LTD.

XX Chase D, Elborough K, Fentem PA, Slabas AR, White A;

XX WPI; 1996-105914/11.

XX N-PSDB; AAQ99304.

XX New isolated rape beta-ketoreductase DNA - used to develop plants

XX with lower or higher oil contents or with altered oil compen.

XX Claim 1; Page 15; 29pp; English.

XX The sequence corresponds to a rape seed beta-ketoreductase encoded
CC by a CDNA insert in plasmid pJRS10.1 in Escherichia coli XL1-Blue.
CC A plastid stroma targeting transit peptide is present. DNA
CC encoding the protein may be inserted in a vector or expression
CC cassette. In sense or antisense orientation for expression in an
CC oilseed plant, e.g. for production of transgenic rape plants with
CC low or modified oil content, diversion of metabolism to alternative
CC storage compounds, e.g. starch, protein or engineered polymers, or
CC production of plants with enhanced oil content. The DNA may also
CC be used as a probe to obtain similar genes from other plants. The
CC transit peptide may be used to direct other proteins to seed
CC plastids.

XX Sequence 315 AA;

Query Match 13.78; Score 222.5; DB 17; Length 315;

Best Local Similarity 25.7%; Pred. No. 7.9e-15; Matches 75; Conservative 50; Mismatches 108; Indels 59; Gaps 9;

QY 2 AAP-MNGQVCVVGASRGIGRGIALQCKAGATYIT-GRILDRLRVAAQEAQSLGQCV 59

DB 66 avprvpsvpyvvgasrgigraiaalslgkagckvlyvnyarsakeaeesvskjeayggqal 125

QY 60 PIVCDSSQSESEVRLTFQVDREOQGRDLVDVNNAYAGVQTLTNKAFMETPASKMDDI 119

DB 126 tfggdvskeadveamkta-idawgtldvynn--agi-----trdclllrmkksqwdv 177

QY 120 NNVGILRGHYFCVYGARLMPAGGLIVISS-PGSLQYMFNVPYGVGAACDLAADCA 178

DB 178 idnltyvflctgaatkmmkkrgrllnlasvvgllgnlganyaaakagvgifsktaa 237

QY 179 HELRRHGVSCVSLMPGIQVTEL---LKEHMAKEVLDPPVLKQKSAFSSAETTELSGKC 235

DB 238 regasrnlnvncpgfiasdmaklgedmek----- 269

QY 236 VVALATDPNLTSLSGVLPSCDLARRRGLRDVGRPVODYLSLSSVLSHVS 287

DB 270 -----kllgtlplgrygqpedvag--lveflalspaasyiltg 304

Search completed: July 31, 2002, 15:08:52
Job time: 231 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:09:41 ; Search time 29.98 Seconds

(without alignments)
1159.643 Million cell updates/sec

Title: US-10-006-163-1

Perfect score: 313
Sequence: 1 MAAPMNGCVTVTGASRGIC.....YLPSELRVPRKMTATYSKF 313

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size: 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

```
1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	313	100.0	313	20	AA127004 Human short-chain
2	240	76.7	313	21	AA127004 Human ORF2404
3	240	76.7	313	21	AA127004 Human ORF2404
4	144	46.0	162	22	AA127004 Human ORF2404
5	65	20.8	303	22	AA127004 Human ORF2404
6	49	15.7	68	22	AA127004 Human ORF2404
7	49	15.7	122	22	AA127004 Human ORF2404
8	40	12.8	72	22	AA127004 Human ORF2404
9	40	12.8	127	22	AA127004 Human ORF2404
10	28	8.9	28	22	AA127004 Human ORF2404
11	24	7.7	24	22	AA127004 Human ORF2404

12	24	7.7	127	22	AA127004
13	24	7.7	127	22	AA127004
14	20	6.4	33	22	AA127004
15	18	5.8	18	22	AA127004
16	18	5.8	23	22	AA127004
17	15	4.8	15	22	AA127004
18	10	3.2	241	22	AA127004
19	10	3.2	244	21	AA127004
20	10	3.2	246	21	AA127004
21	10	3.2	246	21	AA127004
22	10	3.2	246	21	AA127004
23	10	3.2	246	21	AA127004
24	10	3.2	246	21	AA127004
25	10	3.2	246	21	AA127004
26	10	3.2	246	21	AA127004
27	10	3.2	246	21	AA127004
28	10	3.2	255	21	AA127004
29	10	3.2	258	21	AA127004
30	10	3.2	261	21	AA127004
31	10	3.2	286	22	AA127004
32	10	3.2	315	17	AA127004
33	10	3.2	315	17	AA127004
34	9	2.9	241	22	AA127004
35	9	2.9	241	22	AA127004
36	9	2.9	247	21	AA127004
37	9	2.9	247	21	AA127004
38	9	2.9	256	17	AA127004
39	9	2.9	280	21	AA127004
40	9	2.9	321	22	AA127004
41	8	2.6	37	22	AA127004
42	8	2.6	89	22	AA127004
43	8	2.6	120	22	AA127004
44	8	2.6	120	22	AA127004
45	8	2.6	169	22	AA127004

ALIGNMENTS

RESULT 1	AA127004	standard; Protein: 313 AA.
ID	AA127004	standard; Protein: 313 AA.
AC	AA127004	
XX		
DT	20-SEP-1999	(first entry)
XX		
DE	Human short-chain dehydrogenase (HSCD) enzyme.	
XX		
KW	Short-chain dehydrogenase; HSCD; enzyme; pyruvate; coenzyme A; human;	
KW	mitochondria; immune disorder; cancer; leukemia; adenocarcinoma;	
KW	lymphoma; breast; lung; testis; prostate; brain; Addison's disease;	
KW	acquired immune deficiency syndrome; asthma; anemia; Crohn's disease;	
KW	Graves disease; AIDS; gene therapy.	
OS	Homo sapiens.	
XX		
PN	US5928923-A.	
XX		
PD	27-JUL-1999.	
XX		
PF	05-FEB-1998;	98US-0019216.
XX		
PR	05-FEB-1998;	98US-0019216.
XX		
PA	(INCY-) INCYTE PHARM INC.	
XX		
PI	Corley NC, Lal P;	
XX		
DR	WPI; 1999-429503/36.	
XX		
DR	N-PSDB; AAX86772.	
XX		
PT	Nucleic acids encoding human short chain dehydrogenase enzymes	

Human colon cancer
Human gene 22-enco
Human gene 22-enco
Human gene 22-enco
Human gene 22-enco
Human gene 22-enco
Human P24 protein-
S. spina protein
Amino acid sequenc
E. coli cellular p
Staphylococcus aur
Staphylococcus aur
Staphylococcus aur
Staphylococcus aur
Staphylococcus aur
S. aureus NADPH-de
Amino acid sequenc
ORF115 protein inv
Bacillus D-arabinol
Drosophila melanog
Pseudomonas aerugi
Rape leaf beta-ket
Rape seed beta-ket
Putative P. abyssi
Candida magnoliae
3-ketoacyl-ACP red
Pseudomonas aerugi
Glucanase:NADP+-5-
Streptomyces collii
Drosophila melanog
Human gene 22-enco
Novel human secret
Peptide #8053 enco
Human bone marrow
Human polypeptide

PT useful for the diagnosis, prevention and treatment of disorders
 PT associated with abnormal expression of the enzyme, such as Immune
 PT disorders and cancer

PS Claim 1; Fig 1A-D; 27pp; English.

XX This represents a human short-chain dehydrogenase (HSD) enzyme. The
 CC dehydrogenase enzymes catalyze an irreversible reaction between pyruvate
 CC and coenzyme A, to form CO2 and the intermediate CoA, in mitochondria.
 CC Host cells containing vectors comprising the HSD nucleic acid may be
 CC used to produce the HSD enzyme, according to standard recombinant DNA
 CC methodology. The enzyme may then be used as an antigen in the production
 CC of antibodies or in assays to identify antagonists of HSD activity.
 CC These antagonists may then be used to treat disorders associated with
 CC inappropriate expression, or over activity of HSD such as Immune
 CC disorders and cancers (the antagonist interferes with the reaction
 CC between pyruvate and coenzyme A). For example, the antagonists may be
 CC used to treat leukemia, lymphomas, adenocarcinomas and cancers of the
 CC breast, lung, testis, prostate and brain, Addison's disease, acquired
 CC immune deficiency syndrome (AIDS), asthma, anemia, Crohn's disease and
 CC Graves disease. The nucleic acids and antisense sequences can be used
 CC in gene therapy.

SO Sequence 313 AA:

Query Match 100.0%; Score 313; DB 20; Length 313;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAAPMNGOVGVWVGASNGISGRTGIALOLCKAGATVYITGRHLDTLRVVAQNAOSLGQCV 60
 DB 1 maapmngovcvvgasngisgrtgrialolckagatvyltgrhlldlrvaagaqsl99gcvp 60
 OY 61 VCDSSGSESEVRLFEQVDREOQGRDLVLYNNAYAGVOTILTRNKAFWETPASMDDIN 120
 DB 61 vcdssgsesevrlfegvdreoggrldvlynnayagvotlltrnkafwetpsamddin 120
 OY 121 NVGLRGHYFCSYVGARLWPAVAGGGLYVSSPSLQYMFVNPYGVGKAACDKLAADCAHE 180
 DB 121 nvglrghyfcsyvgarlwpavagqglvyspslqymfvnpvgvgaacklaadcahe 180
 OY 181 LRRHGVCVSLMFGIVOTELKREHMAKEEVLQDPVLKQFSAFSAETTELSCGVVLA 240
 DB 181 lrrhgvcsvslmfgivotellkrehmakeevlqdpvlkqfsafsaetelscgvvla 240
 OY 241 TDENITLSLGSKVLPSCDLARRRGLRDVGRPVVDYLSLSSVLSHVSGLWLAATYLPFLR 300
 DB 241 tdenitlslgskvlpscdlarrgrldvgrpvvdyllsssvlshvsglwlaatylpflr 300
 OY 301 VPKWITLALYTSKF 313
 DB 301 vpkwitalytskfi 313

RESULT 2

AA042640 ID AAB42640 standard; Protein: 313 AA.

AC AAB42640;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2404 polypeptide sequence SEQ ID NO:4808.

KX Human: open reading frame; ORFX; detection; cytosolic; hepatotropic;
 KM vlinereary; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
 KM anticonvulsant; osteoplastic; antitachycardic; immunosuppressant; cardiant;
 KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KM hypotensive; dermatological; immunosuppressive; antinflammatory;
 KM antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KM antidiabetic; gene therapy; cancer; proliferative disorder; hypertension;
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;

KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KM cholesterol ester storage; systemic lupus erythematosus; infection;
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KM bone damage; cartilage damage; antinflammatory disease; coagulation;
 KM thrombosis; contraceptive.

OS Homo sapiens.

PN MO200058473-A2.

XX 05-OCT-2000.

PF 31-MAR-2000; 2000MO-US08621.

XX 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.

PR 05-APR-1999; 99US-0127728.

PR 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach M;

DR WPI: 2000-602362/57.

DR N-PSDB; AACT6849.

XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 3990; 5507pp; English.

CC AACT74446 to AACT77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytosolic; hepatotropic; vulnereary;
 CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
 CC osteoplastic; anticonvulsant; antitachycardic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antidiabetic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

SO Sequence 313 AA:

Query Match 76.7%; Score 240; DB 21; Length 313;
 Best Local Similarity 100.0%; Pred. No. 2,1e-234;
 Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 74 LFEQVDBEOGRLDVLVNNAYAGVOTILTRNKAFWETPASMDDINNVGLRGHYFCSY 133
 DB 74 lfegvdbreggrldvlynnayagvotlltrnkafwetpsamddinnvylrghyfcsvy 133
 OY 134 GARLWPAVAGGGLYVSSPSLQYMFVNPYGVGKAACDKLAADCAHELRHGVCVSLMP 193
 DB 134 garlwpavagqglvyspslqymfvnpvgvgaacklaadcahe lrrhgvcsvslmp 193
 OY 194 GIVOTELKREHMAKEEVLQDPVLKQFSAFSAETTELSCGVVLAATDPNITLSLGSKVL 253
 DB 194 givotellkrehmakeevlqdpvlkqfsafsaetelscgvvlaatdpnitslgskvl 253

XX Sequence 162 AA;
SO
Query Match 46.0%; Score 144; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.9e-137;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 170 CDLADCAHELRHGVSCVSLMPGIVQTELEKHNAAKEEVLQDPVYKOFKSAFSSAETT 229
19 cdklaadcahelelrhgvscvslmpgivqteellekhnakeevlqdpvlykqfksafssaett 78
QY 230 ELGKCVVALATPNTLSIGKVLPSGCDLARRVGLRDVGRPVQDLSLSVSHSGSG 289
DB 79 elgkcvvalatpntlsigkvlpsgcdlarryglrdvgrpdyllsbsvlshvsglq 138
QY 290 WLASYPSFLRPVKWITATYTSKF 313
DB 139 wlasypsfirvpkwitalytskf 162
RESULT 5
AAB93802
ID AAB95802 standard; Protein; 303 AA.
AC AAB95802;
XX
DT 26-JUN-2001 (first entry)
DE Human protein sequence SEQ ID NO:18783.
XX
KM Human; primer: detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
EP1074617-A2.
PN
XX
07-FEB-2001.
PD
XX
28-JUL-2000; 2000EP-0116126.
PF
XX
29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX
(HELI-) HELIX RES INST.
XX
Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
WPI: 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX
Claim 8; SEQ ID 18783; 2537pp + CD ROM; English.
XX
XX
The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification, where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC of an oligonucleotide comprising at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SO Sequence 303 AA;
Query Match 20.8%; Score 65; DB 22; Length 303;
Best Local Similarity 100.0%; Pred. No. 3.2e-57;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 MAAPMNGVCVVTGASRGIGRGIALQLCKAGATVYITGRHLDTLRVVAQEAQSLGQCVP 60
DB 1 maapmngvcvvtgasaargigrgialqlckagatvyitgrhldtlrvvaqeaqslgqcvp 60
QY 61 VVCDS 65
DB 61 vvcds 65
RESULT 6
AAG73485
ID AAG73485 standard; Protein; 68 AA.
AC AAG73485;
XX
DT 10-AUG-2001 (first entry)
DE Human gene 22-encoded secreted protein fragment, SEQ ID NO:260.
XX
XX
Human; secreted protein; proliferative disorder; cancer; chromosome 14;
KM foetal abnormality; developmental abnormality; haematopoietic disorder;
KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KM inflammation; allergy; neurological disorder; Alzheimer's disease;
KM Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KM skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KM cardiovascular disorder; angioyenic disorder; kidney disorder;
KM gastrointestinal disorder; pregnancy-related disorder; tumour;
KM endocrine disorder; infection; wound healing; vulnerability;
KM cell culture; chemotaxis; food additive;
KM binding partner identification.
XX
XX
Homo sapiens.
OS
XX
WO200134628-A1.
PN
XX
17-MAY-2001.
PD
XX
08-NOV-2000; 2000WO-US30653.
PF
XX
12-NOV-1999; 99US-0164735.
PR 27-JUL-2000; 2000US-0221193.
PR
XX
(HMDA-) HMDA GENOME SCI INC.
XX
Ruben SM, Komatsoulis GA, Barse CE, NI J, Moore PA;
PI
XX
WPI: 2001-329066/34.
XX
DR Nucleic acids encoding 35 human secreted polypeptides, useful for
XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's
XX disease and diabetic retinopathy -
XX
XX
Disclosure; Page 47; 604pp; English.
XX
XX
AAH32522-AAH32627 represent cDNAs corresponding to 35 human secreted

CC protein genes, and AAG7346-AAG7348 represent the proteins they encode.
 CC AAG73449-AAG73519 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 52 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiotensin-related disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.
 XX Sequence 68 AA:

Query Match 15.7%; Score 49; DB 22; Length 68;
 Best Local Similarity 100.0%; Pred. No. 1.4e-41;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 170 CDKLAADCAHELRHGVCSVSLMPGIVGTETLLKEHMAKEEVLQDPVLKQ 218
 Db 1 cdklaadcahehrrhgvscsvslmpgivgtetllkehmakeevlqdpvlkq 49

RESULT 7
 AAG73483
 ID AAG73483 standard; Protein; 122 AA.
 AC AAG73483;
 XX 10-AUG-2001 (first entry)
 XX Human gene 22-encoded secreted protein fragment, SEQ ID NO:258.
 XX Human; secreted protein; proliferative disorder; cancer; chromosome 14;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiotensin-related disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder; tumour;
 KW endocrine disorder; infection; wound healing; vulnery; food additive;
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification.

OS Homo sapiens.
 XX
 XX WO200134628-A1.
 XX 17-MAY-2001.
 XX 08-NOV-2000; 2000WO-US30653.
 XX 12-NOV-1999; 99US-0164735.
 XX 27-JUL-2000; 2000US-0221193.
 PR

XX (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Komatsoulis GA, Birse CE, N1 J, Moore PA;
 XX WPI: 2001-329066/34.
 XX Nucleic acids encoding 35 human secreted polypeptides, useful for
 XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 XX disease and diabetic retinopathy -
 XX Disclosure; Page 47; 604pp; English.

CC AAH32522-AAH32627 represent cDNAs corresponding to 35 human secreted
 CC protein genes, and AAG73449-AAG73448 represent the proteins they encode.
 CC AAG73449-AAG73519 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 52 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiotensin-related disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.
 XX Sequence 122 AA:

Query Match 15.7%; Score 49; DB 22; Length 122;
 Best Local Similarity 100.0%; Pred. No. 2.3e-41;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 170 CDKLAADCAHELRHGVCSVSLMPGIVGTETLLKEHMAKEEVLQDPVLKQ 218
 Db 12 cdklaadcahehrrhgvscsvslmpgivgtetllkehmakeevlqdpvlkq 60

RESULT 8
 AAG73481
 ID AAG73481 standard; Protein; 72 AA.
 AC AAG73481;
 XX 10-AUG-2001 (first entry)
 XX Human gene 22-encoded secreted protein fragment, SEQ ID NO:256.
 XX Human; secreted protein; proliferative disorder; cancer; chromosome 14;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiotensin-related disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder; tumour;

KW endocrine disorder; infection; wound healing; vulnary;
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification.
 OS Homo sapiens.
 PN WO200134628-A1.
 XX 17-MAY-2001.
 PD 08-NOV-2000; 2000MO-US30653.
 PF 12-NOV-1999; 99US-0164735.
 PR 27-JUL-2000; 2000US-0221193.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Ruben SM, Komatsoulis GA, Birse CE, NI J, Moore PA;
 PI WPI; 2001-329066/34.
 DR Nucleic acids encoding 35 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 PS Disclosure; Page 47; 604pp; English.
 XX AAH3252-AAH3267 represent cDNAs corresponding to 35 human secreted
 CC protein genes, and AAG73446-AAG73448 represent the proteins they encode.
 CC AAG73449-AAG73519 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 52 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.
 XX Sequence 72 AA:

Query Match 12.8%; Score 40; DB 22; Length 72;
 Best Local Similarity 100.0%; Pred. No. 1.9e-32;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 242.DPILSLSGKVPSCDLARRYGARDVGRVQDYLSSV 281
 1 dpilslsgkvpscdlarrylrdvgrvpqdyllsslv 40

RESULT 9
 AAG73479
 ID AAG73479 standard; Protein; 127 AA.
 XX

AC AAG73479;
 XX 10-AUG-2001 (first entry)
 DE Human gene 22-encoded secreted protein fragment, SEQ ID NO:254.
 XX
 KW Human; secreted protein; proliferative disorder; cancer; chromosome 14;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder; tumour;
 KW endocrine disorder; infection; wound healing; vulnary;
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification.
 KM
 XX Homo sapiens.
 OS
 PN WO200134628-A1.
 XX 17-MAY-2001.
 PD 08-NOV-2000; 2000MO-US30653.
 PF 12-NOV-1999; 99US-0164735.
 PR 27-JUL-2000; 2000US-0221193.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Ruben SM, Komatsoulis GA, Birse CE, NI J, Moore PA;
 PI WPI; 2001-329066/34.
 DR Nucleic acids encoding 35 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 PS Disclosure; Page 47; 604pp; English.
 XX AAH3252-AAH3267 represent cDNAs corresponding to 35 human secreted
 CC protein genes, and AAG73446-AAG73448 represent the proteins they encode.
 CC AAG73449-AAG73519 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 52 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.
 XX Sequence 127 AA:

Query Match 12.8%; Score 40; DB 22; Length 127;
 Best Local Similarity 100.0%; Pred. No. 3e-32;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 DPNIISLSGKVLPSCDLARRYGLRDVGRPVODYLSLSV 281
 |||||
 DB 56 dpnllslsgkvlpscdlarryglrdvgrpvdyalslsv 95

RESULT 10
 AAG73482
 ID AAG73482 standard; peptide: 28 AA.
 AC AAG73482;
 DT 10-AUG-2001 (first entry)
 XX Human gene 22-encoded secreted protein fragment, SEQ ID NO:257.
 XX Human; secreted protein; proliferative disorder; cancer; chromosome 14;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angioinfectious disorder; kidney disorder;
 KW gastrointestinal disorder; infection; pregnancy-related disorder; tumour;
 KW endocrine disorder; infection; wound healing; vulnary;
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification.
 XX Homo sapiens.
 OS
 XX WO200134628-A1.
 PN 17-MAY-2001.
 PD
 XX 08-NOV-2000; 2000WO-US30653.
 PF
 XX 12-NOV-1999; 99US-0164735.
 PR 27-JUL-2000; 2000US-0221193.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Komatsoulis GA, Birse CE, NI J, Moore PA;
 XX WPI; 2001-329066/34.
 DR
 XX Nucleic acids encoding 35 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 PT
 XX
 PS Disclosure: Page 47; 604pp; English.
 XX
 AA32522-AAH32627 represent cDNAs corresponding to 35 human secreted
 CC protein genes, and AAG73448 represent the proteins they encode.
 CC AAG73449-AAG73519 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 32 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine

CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.
 CC
 XX Sequence 28 AA;
 XX

Query Match 8.9%; Score 28; DB 22; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.2e-20;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 DPNIISLSGKVLPSCDLARRYGLRDVGR 269
 |||||
 DB 1 dpnllslsgkvlpscdlarryglrdvgr 28

RESULT 11
 AAG73488
 ID AAG73488 standard; peptide: 24 AA.
 AC AAG73488;
 DT 10-AUG-2001 (first entry)
 XX
 XX Human gene 22-encoded secreted protein fragment, SEQ ID NO:263.
 DE
 XX Human; secreted protein; proliferative disorder; cancer; chromosome 14;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angioinfectious disorder; kidney disorder;
 KW gastrointestinal disorder; infection; pregnancy-related disorder; tumour;
 KW endocrine disorder; infection; wound healing; vulnary;
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification.
 KW
 XX Homo sapiens.
 OS
 XX WO200134628-A1.
 PN 17-MAY-2001.
 PD
 XX 08-NOV-2000; 2000WO-US30653.
 PF
 XX 12-NOV-1999; 99US-0164735.
 PR 27-JUL-2000; 2000US-0221193.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Komatsoulis GA, Birse CE, NI J, Moore PA;
 XX WPI; 2001-329066/34.
 DR
 XX Nucleic acids encoding 35 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 PT
 XX
 PS Disclosure: Page 47; 604pp; English.
 XX
 AA32522-AAH32627 represent cDNAs corresponding to 35 human secreted
 CC protein genes, and AAG73448 represent the proteins they encode.
 CC AAG73449-AAG73519 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,

CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the
CC mutations in the new genes. Specific uses are described for each of the
CC 52 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, kidney disorders, asthma,
CC cardiovascular disorders, angiodysplasia, diabetes, atherosclerosis,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infectious. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein fragment referred to in the disclosure of the invention.
SQ Sequence 24 AA;

Query Match 7.7%; Score 24; DB 22; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.1e-16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 218 QFKSAFSAFTTSLSGKCVVALAT 241
Db 1 qfkssaisaeetelsgkcvvalat 24
|||||

RESULT 12

AAG73927
ID AAG73927 standard; Protein; 127 AA.

AC AAG73927;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen protein SEQ ID NO:4691;

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;

KM colorectal carcinoma.

OS Homo sapiens.

XX

PN NC020122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000MO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX DR N-PSDB; AAH33358.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 11; Page 6490-6491; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG7788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in the prevention,
CC therapy and vaccine production. N and P may be used in gene
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
SQ Sequence 127 AA;

Query Match 7.7%; Score 24; DB 22; Length 127;
Best Local Similarity 100.0%; Pred. No. 4.8e-16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 218 QFKSAFSAFTTSLSGKCVVALAT 241
Db 88 qfkssaisaeetelsgkcvvalat 111
|||||

RESULT 13

AAG73486
ID AAG73486 standard; Protein; 127 AA.

AC AAG73486;

DT 10-AUG-2001 (first entry)

DE Human gene 22-encoded secreted protein fragment, SEQ ID NO:261.

KW Human; secreted protein; proliferative disorder; cancer; chromosome 14;

KM foetal abnormality; developmental abnormality; haematopoietic disorder;

KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;

KW inflammation; allergy; neurological disorder; Alzheimer's disease;

KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;

KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;

KW cardiovascular disorder; angiodysplasia; kidney disorder;

KW gastrointestinal disorder; pregnancy-related disorder; tumour;

KW endocrine disorder; infection; wound healing; vulvular;

KW cell culture; chemotaxis; food additive;

KW binding partner identification.

OS Homo sapiens.

XX WO200134628-A1.

XX 17-MAY-2001.

XX 08-NOV-2000; 2000MO-US30653.

XX 12-NOV-1999; 99US-0164735.

XX 27-JUL-2000; 2000US-0221193.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Komatsoulis GA, Birse CE, NI J, Moore PA;

XX WPI; 2001-329066/34.

XX Nucleic acids encoding 35 human secreted polypeptides, useful for

PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PR disease and diabetic retinopathy -
 XX
 PS Disclosure; Page 47; 604pp; English.

XX
 XX AAH32522-AAH32627 represent cDNAs corresponding to 35 human secreted
 CC protein genes, and AAG73449-AAG73448 represent the proteins they encode.
 CC AAG73449-AAG73519 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 52 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.
 XX
 SQ Sequence 127 AA;

Query Match 7.7%; Score 24; DB 22; Length 127;
 Best Local Similarity 100.0%; Pred. No. 4.8e-16;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 218 QFKSAFSAETTELGGKCVVALAT 241
 ||||||||||||||||||||
 Db 88 qfkssafsaetfelsgkcvvalat 111

RESULT 14
 AAG73484
 ID AAG73484 standard; peptide: 33 AA.

XX
 AC AAG73484;
 XX
 DT 10-AUG-2001 (first entry)
 XX

DE Human gene 22-encoded secreted protein fragment, SEQ ID NO:259.

XX Human; secreted protein; proliferative disorder; cancer; chromosome 14;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder; tumour;
 KW endocrine disorder; infection; wound healing; vulnery; food
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification.

XX Homo sapiens.

OS WO200134628-A1.

XX
 XX
 PN
 XX

PD 17-MAY-2001.
 XX
 XX 08-NOV-2000; 2000WO-US30653.
 PE
 XX 12-NOV-1999; 99US-0164735.
 PR 27-JUL-2000; 2000US-0221193.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Komatsoulis GA, Birse CE, NI J, Moore PA;

DR WPI: 2001-329066/34.

PT Nucleic acids encoding 35 human secreted polypeptides, useful for
 PR preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PR disease and diabetic retinopathy -
 XX
 XX

PS Disclosure; Page 47; 604pp; English.

CC AAH32522-AAH32627 represent cDNAs corresponding to 35 human secreted
 CC protein genes, and AAG73449-AAG73448 represent the proteins they encode.
 CC AAG73449-AAG73519 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 52 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.
 XX
 SQ Sequence 33 AA;

Query Match 6.4%; Score 20; DB 22; Length 33;
 Best Local Similarity 100.0%; Pred. No. 1.7e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 199 ELIKEMAKEEVLQDPVLKQ 218
 ||||||||||||||||||||
 Db 1 elikemakeevlqdpvlkq 20

RESULT 15
 AAG73487
 ID AAG73487 standard; peptide: 18 AA.

XX
 AC AAG73487;
 XX
 DT 10-AUG-2001 (first entry)
 XX

DE Human gene 22-encoded secreted protein fragment, SEQ ID NO:262.

XX Human; secreted protein; proliferative disorder; cancer; chromosome 14;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;

Search completed: July 31, 2002, 15:13:33
 Job time: 232 sec

KW Immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiotensin-related disorder; kidney disorder;
 KW endocrine disorder; infection; pregnancy-related disorder; tumour;
 KW cell culture; chemotaxis; food additive; vulvular;
 KW binding partner identification.
 OS Homo sapiens.
 XX
 XX
 PN WO200134628-A1.
 PD 17-MAY-2001.
 XX
 XX
 PF 08-NOV-2000; 2000WO-US30653.
 XX
 PR 12-NOV-1999; 99US-0164735.
 XX 27-JUL-2000; 2000US-0221193.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX
 PI Ruben SM, Komatsoulis GA, Birse CE, NI J, Moore PA;
 DR WPI; 2001-329066/34.
 XX
 XX
 PT Nucleic acids encoding 35 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy.
 PS Disclosure; Page 47; 604pp; English.
 XX
 XX
 CC AAH32522-AAH32627 represent CDNA's corresponding to 35 human secreted
 CC protein genes, and AAG73446-AAG73448 represent the proteins they encode
 CC and their corresponding secreted protein fragments. The genes
 CC treating or ameliorating medical conditions, e.g., by preventing,
 CC amount of the new protein in a sample or by determining the
 CC mutations in the new genes. Specific uses are described for each of the
 CC 52 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC cardiovascular disorders, angiotensin-related disorders, atherosclerosis,
 CC gastrointestinal disorders, sepsis, diabetes, kidney disorders,
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC as a food additive or binding partners, and in chemotaxis, and can be used
 CC antibodies specific for a protein of the invention can be used
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.
 XX
 S0 Sequence 18 AA;

Query Match

Best Local Similarity 5.88; Score 18; DB 22; Length 18;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 FSSAETTESGRCVVALA 240
 DB 1 ISSAETTESGRCVVALA 18
 1 ISSAETTESGRCVVALA 18

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:08:56 ; Search time 30.15 Seconds

(without alignments)
1795,934 Million cell updates/sec

Title: US-10-006-163-1
Sequence: 1623
1 MAPNMGVCVVTGASRGIG.....YLPSFLRVKMTIALYTSKF 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 1 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOZOA:*
12: SP_VIRUS:*
13: SP_VIRIDIA:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEA:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1619	99.8	313	4	Q96LJ7
2	1615	99.5	313	4	Q96B59
3	1609	99.1	313	4	Q96C05
4	1391	85.7	313	11	Q99L04
5	1370	84.4	313	11	Q99L48
6	710	43.7	319	5	Q9N538
7	615	37.9	323	5	Q16764
8	583	35.9	325	5	Q23612
9	347.5	21.4	326	2	Q54118
10	343.5	21.2	302	16	Q9KFR7
11	338.5	20.9	302	16	Q98GEO
12	335.5	20.7	308	2	Q98W72
13	310	19.1	255	16	Q912R7
14	286.5	17.7	248	16	Q92AX2
15	271	16.7	255	16	P95286
16	262	16.1	250	16	P71079

17	247	15.2	249	16	Q9KEB5	Q9KEB5 bacillus ha
18	247	15.2	266	16	Q93906	Q93906 eubacterium
19	243	15.0	266	16	Q9HWN3	Q9HWN3 pseudomonas
20	241	14.8	249	16	Q97DA6	Q97DA6 clostridium
21	239	14.7	261	10	Q9FK50	Q9FK50 arabidopsis
22	234.5	14.4	247	16	Q92AK1	Q92AK1 listeria in
23	234.5	14.4	253	2	Q9K377	Q9K377 streptomyces
24	232	14.3	268	17	Q9HNL6	Q9HNL6 thermoplasma
25	227.5	14.0	270	5	Q9VNF3	Q9VNF3 drosophila
26	227.5	14.0	257	10	Q9S0R4	Q9S0R4 arabidopsis
27	227	14.0	254	5	Q95PA6	Q95PA6 aedes aegypti
28	226.5	14.0	272	10	Q9S0R2	Q9S0R2 arabidopsis
29	224.5	13.8	255	2	Q9APX4	Q9APX4 pseudomonas
30	224.5	13.8	258	16	Q92RTX5	Q92RTX5 rhizobium m
31	224.5	13.8	271	2	Q92336	Q92336 clostridium
32	222.5	13.7	273	2	Q9EF51	Q9EF51 streptomyces
33	222.5	13.7	315	10	Q949M3	Q949M3 brassica na
34	222.5	13.7	320	10	Q93X62	Q93X62 brassica na
35	221	13.6	254	5	Q95PA3	Q95PA3 aedes aegypti
36	221	13.6	254	10	Q949M2	Q949M2 brassica na
37	220.5	13.6	328	10	Q93X67	Q93X67 brassica na
38	220.5	13.6	374	16	Q989M0	Q989M0 rhizobium 1
39	218	13.4	251	16	Q98FG6	Q98FG6 rhizobium 1
40	216.5	13.3	254	2	Q9RH24	Q9RH24 zymomonas m
41	216.5	13.3	278	16	Q9A3M7	Q9A3M7 caulobacter
42	216	13.3	246	16	Q9KA03	Q9KA03 bacillus ha
43	216	13.3	257	16	Q987C7	Q987C7 rhizobium 1
44	214.5	13.2	246	16	Q99QK7	Q99QK7 staphylococcus
45	214.5	13.2	248	16	Q97FV0	Q97FV0 clostridium

ALIGNMENTS

RESULT 1
ID Q96LJ7 PRELIMINARY: PRT: 313 AA.
AC Q96LJ7:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CDNA FLJ25430 F1S, CLONE TST06262.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Suzuki Y., Hata H.,
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura T., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,
RA Nagai K., Isogai T., Sugano S.:
RT "NEPO human cDNA sequencing project."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, AK08159; BAB71694.1;
SQ SEQUENCE 313 AA; 33909 MW; 832F83FA75D931A3 CRC64;

Query Match 99.8%; Score 1619; DB 4; Length 313;
Best Local Similarity 99.7%; Pred. No. 1.2e-133;
Matches 312; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAPNMGVCVVTGASRGIGRGIALQLCKAGATVYITGRHLDLTVVVAQEAQSLGGQCP 60
DB 1 MAPNMGVCVVTGASRGIGRGIALQLCKAGATVYITGRHLDLTVVVAQEAQSLGGQCP 60
QY 61 VVCDSSQSESVPTLEQVVDREOGRIDVLYNNAYAGVOTILTRNKAEKETPASMDDIN 120
DB 61 VVCDSSQSESVPTLEQVVDREOGRIDVLYNNAYAGVOTILTRNKAEKETPASMDDIN 120
QY 121 NVGLRGHRCFSYGARIMVPAGOGGLIVYISSPSGLQYMNVPYGYGKACADLADCAHE 180

```

Db 121 NVGLGHYFCVYGARLWVPAOGGLIYVSSPSGLQYMFENPYVGAACDKLAADCAHE 180
QY 181 LRRHGVSCVSLMPGIYOTELKEHMAKEEVLQDPVLKQFSAFSAETTELSCGCYVALA 240
Db 181 LRRHGVSCVSLMPGIYOTELKEHMAKEEVLQDPVLKQFSAFSAETTELSCGCYVALA 240
QY 241 TDPNIIISLCKVLPSCDLARRYGLRDVGRPVODYLSLSSVSHVSGLGWLASTLPSEFLR 300
Db 241 TDPNIIISLCKVLPSCDLARRYGLRDVGRPVODYLSLSSVSHVSGLGWLASTLPSEFLR 300
QY 301 VPKWIIALYTSKF 313
Db 301 VPKWIIALYTSKF 313

```

```

RESULT 2
096B59 ID 096B59 PRELIMINARY; PRT; 313 AA.
AC 096B59;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE HYPOHETICAL 33.9 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN NCBI_Taxid=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=MELANOMA;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015943; AAH15943.1;
KW Hypothetical protein.
SQ SEQUENCE 313 AA; 33881 MW; ED0D252724E38776 CRC64;

```

```

Query Match
Best Local Similarity 99.58; Score 1615; DB 4; Length 313;
Matches 311; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MAAPMNGVCVVTGASRGIRGIALQCKAGATVYITGRHLDTLRYVAOASLGCCVP 60
Db 1 MAAPMNGOACVVTGASRGIRGIALQCKAGATVYITGRHLDTLRYVAOASLGCCVP 60
QY 61 VVCDSSQSESEVRLFEQVDREOGRDLVLYNNAYAGVOTILNTRNKAFTETPASMDDIN 120
Db 61 VVCDSSQSESEVRLFEQVDREOGRDLVLYNNAYAGVOTILNTRNKAFTETPASMDDIN 120
QY 121 NVGLGHYFCVYGARLWVPAOGGLIYVSSPSGLQYMFENPYVGAACDKLAADCAHE 180
Db 121 NVGLGHYFCVYGARLWVPAOGGLIYVSSPSGLQYMFENPYVGAACDKLAADCAHE 180
QY 181 LRRHGVSCVSLMPGIYOTELKEHMAKEEVLQDPVLKQFSAFSAETTELSCGCYVALA 240
Db 181 LRRHGVSCVSLMPGIYOTELKEHMAKEEVLQDPVLKQFSAFSAETTELSCGCYVALA 240
QY 241 TDPNIIISLCKVLPSCDLARRYGLRDVGRPVODYLSLSSVSHVSGLGWLASTLPSEFLR 300
Db 241 TDPNIIISLCKVLPSCDLARRYGLRDVGRPVODYLSLSSVSHVSGLGWLASTLPSEFLR 300
QY 301 VPKWIIALYTSKF 313
Db 301 VPKWIIALYTSKF 313

```

```

RESULT 3
096C05 ID 096C05 PRELIMINARY; PRT; 313 AA.
AC 096C05;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

```

```

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SIMILAR TO RIKEN CDNA 1110029G07 GENE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN NCBI_Taxid=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREATIC ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014057; AAH14057.1;
KW Oxidoreductase.
SQ SEQUENCE 313 AA; 33925 MW; 37FA02675C4F076 CRC64;

```

```

Query Match
Best Local Similarity 99.18; Score 1609; DB 4; Length 313;
Matches 311; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MAAPMNGVCVVTGASRGIRGIALQCKAGATVYITGRHLDTLRYVAOASLGCCVP 60
Db 1 MAAPMNGOACVVTGASRGIRGIALQCKAGATVYITGRHLDTLRYVAOASLGCCVP 60
QY 61 VVCDSSQSESEVRLFEQVDREOGRDLVLYNNAYAGVOTILNTRNKAFTETPASMDDIN 120
Db 61 VVCDSSQSESEVRLFEQVDREOGRDLVLYNNAYAGVOTILNTRNKAFTETPASMDDIN 120
QY 121 NVGLGHYFCVYGARLWVPAOGGLIYVSSPSGLQYMFENPYVGAACDKLAADCAHE 180
Db 121 NVGLGHYFCVYGARLWVPAOGGLIYVSSPSGLQYMFENPYVGAACDKLAADCAHE 180
QY 181 LRRHGVSCVSLMPGIYOTELKEHMAKEEVLQDPVLKQFSAFSAETTELSCGCYVALA 240
Db 181 LRRHGVSCVSLMPGIYOTELKEHMAKEEVLQDPVLKQFSAFSAETTELSCGCYVALA 240
QY 241 TDPNIIISLCKVLPSCDLARRYGLRDVGRPVODYLSLSSVSHVSGLGWLASTLPSEFLR 300
Db 241 TDPNIIISLCKVLPSCDLARRYGLRDVGRPVODYLSLSSVSHVSGLGWLASTLPSEFLR 300
QY 301 VPKWIIALYTSKF 313
Db 301 VPKWIIALYTSKF 313

```

```

RESULT 4
099L04 ID 099L04 PRELIMINARY; PRT; 313 AA.
AC 099L04;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE RIKEN CDNA 1110029G07 GENE.
GN 1110029G07 GENE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_Taxid=10090;
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY TUMOR. MAP-TGF ALPHA MODEL. 7 MONTHS OLD, GROSS
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL; BC003930; AAH03930.1;
DR MGD; MGI:1915960; 1110029G07Rik.
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
KW Oxidoreductase.
SQ SEQUENCE 313 AA; 34005 MW; 60E05BD7911BDC0C CRC64;

```

Query Match	85.7%	Score 1391	DB 11	Length 313
Best Local Similarity	82.4%	Pred. No. 1.le-113		
Matches 256	Conservative 27	Mismatches 28	Indels 0	Gaps 0
QY	1	MAAPMGQCVVVTASRIGRIGALQOLCKAATVYITGRHLDLTVHVAOEAQSLGQCV	60	
Db	1	MVAPMGQCVVVTASRIGRIGALQOLCKAATVYITGRHLDLTVHVAOEAQSLGQCV	60	
QY	61	VVCSQSESEVRILPEOVYDREOQRLDVLVNNAVAGVOTILTRFKALWEPASMMDDIN	120	
Db	61	VVCSQSESESEVRILPEOVYDREOQRLDVLVNNAVAGVOTILTRFKALWEPASMMDDIN	120	
QY	121	NVGRGHVFCGVYCARLWVPPAGGCIIVYISPSGLQVFNPNVYGVGKACDLADCAHE	180	
Db	121	NVGRGHVFCGVYCARLWVPPAGGCIIVYISPSGLQVFNPNVYGVGKACDLADCAHE	180	
QY	181	LRRHGVSCVSLMPGIVQTELLKEHAKKEEVLDDPYLQKFSAFSSAETTELSCGCYVALA	240	
Db	181	LRRHGVSCVSLMPGIVQTELLKEHAKKEEVLDDPYLQKFSAFSSAETTELSCGCYVALA	240	
QY	241	TDPNLSISGVVLPSCDLARRRGLRDVDGRPVQDYLSSVLSHVSGWLAATYPSFLR	300	
Db	241	TDPNLSISGVVLPSCDLARRRGLRDVDGRPVQDYLSSVLSHVSGWLAATYPSFLR	300	
QY	301	VPKMIILYTSKF 313		
Db	301	VPKMIILYTSKF 313		
RESULT	5			
ID	Q9D148	PRELIMINARY	PRF	313 AA.
AC	Q9D148			
DT	01-JUN-2001	(TREMBlrel. 17, Created)		
DT	01-JUN-2001	(TREMBlrel. 17, Last sequence update)		
DT	01-DEC-2001	(TREMBlrel. 19, Last annotation update)		
DE	1110029G07R1K	PROTEIN.		
GN	1110029G07R1K			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;			
OX	NCBI_TaxID=10090;			
XP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=EMBRYO;			
RC	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shihagawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi S., Fukuda S.,			
RA	Aizawa K., Itawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,			
RA	Sato T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Koichiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,			
RA	Schirnl L.M., Scabuli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Guatichin S., Hill D., Hofmann M., Hume D.A., Kamlaya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,			
RA	Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyokawa K., Wang K.H., Weitz C., Wittaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,			
RA	Hayashizaki Y.,			
RT	"Functional annotation of a full-length mouse cDNA collection."			
RL	Nature 409:655-690(2001).			
DR	EMBL; AK003958; BAB23093.1.			
DR	MGD; MGI:1915960; 1110029G07R1K.			
SO	SEQUENCE 313 AA; 34031 MW; 808E8ECD7910EDC03 CRC64;			

Matches	256:	Conservative	27:	Mismatches	30:	Indels	0:	Gaps	0:
QY	1	MAAPMNGVCVVTGASNRIGRGIALQJLCAGATYITGRHLDTLRVVAQEAQSLGGQCV	60						
Db	1	MAAPMNGVCVVTGASNRIGRGIALQJLCAGATYITGRHLDTLRVVAQEAQSLGGQCV	60						
QY	61	VVCCSSQSESEFRITLFEQYDREQGRDLVLYNNATAGVTTILNTRKKAWEHPASKMDIN	120						
Db	61	VVCCSSQSESEKSLFEQYDREQGRDLVLYNNATAGVTTILNTRKKAWEHPASKMDIN	120						
QY	121	NVGLRGHFCGVYAGRLVAPVPGOGILIVIVSSPSLOYMENTPYGVGKAACKLADCAHE	180						
Db	121	NVGLRGHILCVYAGRLVAPVPGOGILIVIVSSPSLOYMENTPYGVGKAACRLADCAHE	180						
QY	181	LRRRGVSCVSLMPGIVQVTELLRKHAKKEEVLQDPYLQFKSAFSSAETTELSGKCVALA	240						
Db	181	LRRRGVSVSLMPRLVQVEMVKEFMAKEDPEDPELFKFMKPDFSSAESPEMSGKCVALA	240						
QY	241	TDPNLSLSGCVLPSCLDARRRYGLADVGRVQDYPLSLSSVLSHVSGLGLASYPFLR	300						
Db	241	TDPNLTNLNSGCVLPSCDLARRRYGLKDIDGRPYKDYFSLGVALSQVSLGWLNSFLPGLR	300						
QY	301	VPKMIATLYTSKF 313							
Db	301	VPKLVVTLTKSKF 313							
RESULT	6								
Q9N538		PRELIMINARY:	PRT:	319 AA.					
AC	Q9N538								
DT	01-OCT-2000 (TREMBLrel. 15, Created)								
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)								
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)								
DE	HYPOTHEMETICAL 35.1 KDA PROTEIN.								
OS	Y32H12A.3								
GN	Caenorhabditis elegans.								
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;								
OC	Rhabditidae; Peloderinae; Caenorhabditis.								
OX	NCBI_Taxid=6239;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN-BRISTOL N2;								
RX	MEDLINE#99069613; PubMed=9851916;								
RA	None;								
RT	"Genome sequence of the nematode C. elegans: a platform for								
RT	investigating biology. The C. elegans Sequencing Consortium."								
RL	Science 282:2012-2018(1998).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN-BRISTOL N2;								
RC	Holmes A.; Elliot G.; Cloud J.;								
RA	"The sequence of C. elegans cosmid Y32H12A."								
RT	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.								
RL	[3]								
RN	SEQUENCE FROM N.A.								
RP	STRAIN-BRISTOL N2;								
RC	Waterston R.;								
RA	"Direct Submission."								
RT	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.								
CC	-1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES								
CC	(SDR) FAMILY.								
DR	EMBL: AC006733; AAF60486.1; -								
DR	HSSP: P50162; IAE1.								
DR	InterPro: IPR002198; ADH_short.								
DR	PRINTS: PR00080; SDRFAMILY.								
DR	PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.								
KW	Hypothetical protein; Oxidoreductase.								
SO	SEQUENCE 319 AA; 35067 MW; FCB5DDAE07EE734E CRC64;								
Query Match	43.7%;	Score 710;	DB 5;	Length 319;					
Best Local Similarity	46.8%;	Pred. No. 5e-54;							

Matches 146; Conservative 51; Mismatches 99; Indels 16; Gaps 4;

OY 7 GOVCVVTGASRGIGIALQLOCKAGATYITGRH-----LDTLRVVAOENOSLGG 56
DB 5 GOIATVTSASRGIGIALQLOGEAGATYITGRKEESLNSKVGSLGEATADEITKNG 64
OY 57 QCVPVCDSSOESSEVRLTFEVDREOQGLDVLVNNAVAGVOTILNFRKAFMETPSMW 116
DB 65 KGIAFVDHOMEEKNEFEVEKEHOGQDILVNNATOGVTAISENKKPFTETDPYV 124
OY 117 DOINNVGLRGHFCFSYVGARLMPAGGGLIYVISPGLQYMFVNPYVGKAAACDKLAAD 176
DB 125 DTINNVGLRHHYFCTVVAARLMTARNKGLIYVSSGGRLRFLFNVAAYGQALDRMSAD 184
OY 177 CAHLRRHGVSCVSLMPGIYOTELKHEMAKEVLODPVYKOFKSAFSSAETELSGKV 236
DB 185 TAVELRRKNCVSVIMGAVATELVDMFDENGKPRPEIKN AEVANGSETVYPCRAV 243
OY 237 VALATDPNIIISLGRVLPSCDLARRYGLRDVGRPVQDYLSLSVLSHVGSLG--LASV 294
DB 244 VSLASDRRRDKTGRIITLEDIGKEVGFVDIDGLRPNLMSVSLTNG---LGNNTTANF 300
OY 295 LPSFLRYPKMTI 306
DB 301 VPTWKLPGWLV 312

RESULT 7
ID 016764 PRELIMINARY: PRT; 323 AA.
AC 016764;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE HYPOTHETICAL 35.8 KDA PROTEIN.
GN F59E11.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae.
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
ON (1)
RA SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Bradshaw H.;
RT "The sequence of C. elegans cosmid r59E11.";
RN Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
RP [3]
RC SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL; AF016685; AAG24139.1; -
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short.1.
KW Hypothetical protein; Oxidoreductase.
SQ SEQUENCE 323 AA; 35824 MW; 4FE24D30ED39CAE3 CRC64;

Query Match 37.9%; Score 615; DB 5; Length 323;
Best Local Similarity 42.1%; Pred. No. 1,1e-45;
Matches 138; Conservative 56; Mismatches 112; Indels 20; Gaps 5;

OY 1 MAAPNGOVVVTGASRGIGIALQLOCKAGATYITGRH-----LDTLRVVAOENOSLGG 50
DB 1 MGVLDDOVALVYTGASRGIGIALQLOGEAGATYITGRBELSDNFRGLPSIDYAKE 60
OY 51 AOSLGOCVPCVCDSSOESSEVRLTFEVDREOQGLDVLVNNAVAGVOTILNFRKAFME 110
DB 61 ITRSGKGIALYVDHNSMTEVKFLFEKIKDEBEKDLVNVNNSLCKATEMIGKTFED 120
OY 111 TPASMDINNVGLRGHFCFSYVGARLMPAGGGLIYVISPGLQYMFVNPYVGKAAAC 170
DB 121 ODPSFMDINGVGLRHHYFCTVVAARLMTARNKGLIYVSSGGRLRFLFNVAAYGQALDRMSAD 180
OY 171 DKLADCAHLRRHGVSCVSLMPGIYOTELKHEM---AKEVLODPVYKOFKSAFSSAE 227
DB 181 ARMSTDAVELNPNVNCVTVLIPQVETETANRTITIDAKMKIENPELE---FTKGE 236
OY 228 TTELSGKCVVALATDPNIIISLGRVLPSCDLARRYGLRDVGRPV--QDYLSLSVLSHV 285
DB 237 STEYTGKALRLAMDPEKLLKSKGTFTEDLAQVDFSDKHGAGMEPQNTRSINTILG-T 295
OY 286 SGLGWLASYLPSFLRYPKMTIALYTSKF 313
DB 296 MKEEVAKITPQIKLPKVMVIMOSVNR 323

RESULT 8
ID 023612 PRELIMINARY: PRT; 325 AA.
AC 023612;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE HYPOTHETICAL 35.9 KDA PROTEIN.
GN ZK816.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae.
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
ON (1)
RA SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Nhan M., Leimbac D.;
RT "The sequence of C. elegans cosmid ZK816.";
RN Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RP [3]
RC SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL; U41018; AAA82327.1; -
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short.1.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Hypothetical protein; Oxidoreductase.
SQ SEQUENCE 325 AA; 35919 MW; D68C60E9105AE177 CRC64;

Query Match 35.9%; Score 583; DB 5; Length 325;
Best Local Similarity 39.2%; Pred. No. 6.8e-43;
Matches 130; Conservative 55; Mismatches 111; Indels 36; Gaps 6;

```

Db      3 LASKIAIYVAGSRGCGRGLKGAVALQALAGCTLYITGRAPSKTSSSELYTFLTGTA 62
Qy      50 EAGSAGGOCVPVYCDSSOSEVFTLEVOYDREOQGRDLVYNNAYGVOITLITRNKA 109
Db      63 ECRKRGIGCHVRYVSHNDEVEKEFEDEVASETDOLDILVNNAFSAVTKCGSGDTRKFF 122
Qy      110 EPPASMDIDNNVGLRGHFCFSCYVGRALNVPAG-OGILYIVSSPSGLQYFENPYGVKA 168
Db      123 EDPDELMDIDNNVGLRNQYCSYVGRIRKMKMGLYINISLIGIMLYFTVAYGVKMA 182
Qy      169 ACCKLAADCAHLELRHGVSCVSLMPGIVOTELLKEHM-----AKEEVLADPVLKQFK 220
Db      183 ALDRKSSDMAOGLDQDTGIVISLMPASVATELITNMTESAGSGATENKM-----233
Qy      221 SAFSSAETTELSCGVVALATDPNIIISLGVLPSCDLARRVGLRDVGRPVDDYLSLS 280
Db      224 --FLNGESTETCYGKAVVAIAADPKKYMGSSTLITDMGNYVYTDIDGRIPTNMQLRG 291
Qy      281 VLSHVSGLWKLASYLPSFLRVKRIITATLSK 312
Db      292 ILS-LAGYSHMAGMCPWVNLPGMAITLWQNK 322

```

RESULT 9

```

ID      054118      PRELIMINARY;      PRT;      326 AA.
AC      054118;
DT      01-JUN-1998 (TREMblrel. 06, Created)
DT      01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT      01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE      SC10A5.27C PROTEIN.
GN      SC10A5.27C.
OS      Streptomyces coelicolor.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
CX      Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
RN      NCBI_Taxid-1902;
[1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-A3(2);
RA      Murphy L., Harris D.;
RL      Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP      SEQUENCE FROM N.A.
RC      STRAIN-A3(2);
RA      Parthali J., Barrell B.G., Rajandream M.A.;
RL      Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP      SEQUENCE FROM N.A.
RC      STRAIN-A3(2);
RA      MEDLINE-97000351; PubMed-8843436;
RA      Redenbach M., Kleser H.M., Denaplatte D., Eichner A., Cullum J.,
RA      Klashi H., Hopwood D.A.;
RA      "A set of ordered cosmid and a detailed genetic and physical map for
RT      the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL      Mol. Microbiol. 21:77-96(1996).
CC      -1- STRAILITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES.
CC      (SDR) FAMILY.
DR      EMBL; AL021529; CAAL6459.1; -.
DR      InterPro; IPR002198; ADH_short.
DR      Pfam; PF00106; adh_short; 1.
KW      Oxidoreductase.
SQ      SEQUENCE 326 AA; 35276 MW; CB8F3D5B218B0EB9 CRC64;

```

```

Query Match      21.4%; Score 347.5; DB 2; Length 326;
Best Local Similarity 33.6%; Pred. No. 2.9e-22;
Matches 99; Conservative 55; Mismatches 102; Indels 39; Gaps 11;
Qy      4 PANGOVCVTVGASRGIGRGIALQLCRAGATVYIT-----GRHLDLRLVNAQEAQ 52
Db      22 PLAGRIATVAGATRGAGRAQAVELGRAGATVYVYGTTRARASEVGTETIETAEVLTV 81

```

```

Qy      53 SLGGOCVPVYCDSSOSEVFTLEVOYDREOQGRDLVYNNAYGVOITLITNR--NKAFFE 110
Db      82 AAGGTGIAVFTDLDEAQAVALVERIDREYE-RDLIDVNDLMGG-EHLATSVFGKSM 139
Qy      111 TPASMDIDNNVGLRGHFCFSCYVGRALNVPAGOGILYIVISSP-----GSLQYMF 159
Db      140 TPLADGLRIELGARSH-----VITALLLP-----LLIRSAFPHVEYTDGTASHNRNRYE 191
Qy      160 NPEYGVKAACDLAADCALHLELRHGVSCVSLMPGIVOTELLKEHM-AKEEVLADPVLKQ 218
Db      192 NIYDLAKNAPILRARGLAQELAEYEGTAVAVSPGLRSPOMLSHFGVSEBNRDAIAOE 251
Qy      219 FKAFSSAETTELSCGVVALATDPN-ILSLSGVLPSCDLARRVGLRDVGR-RP 271
Db      252 --PTFAIESPHYIARTVAAALADPDAKRNKGSSTSGELARAYGVTDVDSRP 304

```

RESULT 10

```

ID      09KFR7      PRELIMINARY;      PRT;      302 AA.
AC      09KFR7;
DT      01-OCT-2000 (TREMblrel. 15, Created)
DT      01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT      01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE      BH0410 PROTEIN.
GN      BH0410.
OS      Bacillus halodurans.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group;
CX      Bacillus/Staphylococcus group; Bacillus.
RN      NCBI_Taxid-86665;
[1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-C-125 / JCM 9153;
RA      MEDLINE-20512582; PubMed-11058132;
RA      Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA      Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA      Horikoshi K.;
RT      "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT      halodurans and genomic sequence comparison with Bacillus subtilis.";
RL      Nucleic Acids Res. 28:4317-4331(2000).
CC      -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC      (SDR) FAMILY.
DR      EMBL; AP001508; BAB04129.1; -.
KW      Complete proteome; Oxidoreductase.
SQ      SEQUENCE 302 AA; 33422 MW; CE13F300140D55BB CRC64;

```

```

Query Match      21.2%; Score 343.5; DB 16; Length 302;
Best Local Similarity 22.5%; Pred. No. 5.8e-22;
Matches 94; Conservative 57; Mismatches 101; Indels 37; Gaps 10;

```

```

Qy      4 PANGOVCVTVGASRGIGRGIALQLCRAGATVYITGR-----HLDTLRVNAQEAQ 53
Db      3 PLIGKVALVAGATRGAGRGIAVELGAAGATVYVYGTTRERRESEYRNPERIETAEVLVN 62
Qy      54 LGGOCVPVYCDSSOSEVFTLEVOYDREOQGRDLVYNNAYGVOITLITRNKAFFETPA 113
Db      63 AGGTGIAVFPVDHDPQKVEALVQIERE-OGKLDVYVNDIMGG-----ELTEWNA 112
Qy      114 SMDMDIDNNVGLR-----GHYFCVYGARLMPAGOGILYIVSSPSGLQY--MENVLP- 162
Db      113 PVOHSLDDELRLRLAIDHLLTSHFALPFLKS-RGLVEMTD-GTAEYNEHYRQPL 170
Qy      163 -YGVKAACDLAADCALHLELRHGVSCVSLMPGIVOTELLKEHM-AKEEVLADPVLKQFK 220
Db      171 YDLATSVLRAMAGLAQELQPHCTAVALTPEWMNSEIMLDHFETERNWRATIKE-- 228
Qy      221 SAFFSAETTELSCGVVALATDPNIIISLGVLPSCDLARRVGLRDVGR 269
Db      229 PHEVISESPFVGRVAVAAALASDPVSRMNGOSLSGOLAQAYGFTDIDG 277

```

RESULT 11

```

09GEO ID 09GEO0 PRELIMINARY: PRT: 302 AA.
AC 09GEO0;
DR 01-OCT-2001 (TREMBLrel. 18, Created)
DR 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DE 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
GN MLJ372 PROTEIN.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_Taxid=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shlimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL, AP003001; BAB50276.1;
DR InterPro; IPR02198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
KW Complete proteome.
SQ SEQUENCE 302 AA; 32880 MW; EC7255ACAB43C337 CRC64;

Query Match
Best Local Similarity 33.0%; Score 338.5; DB 16; Length 302;
Matches 100; Conservative 53; Mismatches 103; Indels 47; Gaps 12;

OY 5 MNGQCVVVGASRGIGGIALQCKAGATYITGRHL-----DLRVAAQEAQS 54
DB 3 LGGCVVAVAGTNRGAGGIVAGATVYVGRSTRAQSEYARETIEETALVTAN 62
OY 55 GGQCVVAVCDSSQSESERTLFEOVDREOGRDLVNNAAVAGOTILNTNKA-FWETPA 113
DB 63 GGSINAVQADHVAADVRGLIER-KREGRLDILVNDING-----EKLFEWDKP- 112
OY 114 SMQDINNNGLR-----GHYCSYVGARLMPVAGGGLIVVSSPSGL-----YMFNV 161
DB 113 -VVEHNDKGLRLRLAVETHAITSHALPDLRRPGVIVEVDGDAVNRDRSFF 177
OY 162 PYGVGKACDKLADCAHELRHGVSCVSLMPGIYOTELKEHMAKEEV-LQDPVLKQFK 220
DB 171 FYDLAVAVNRMAHAKDKAGATSVSLTGMLESEMLEAFGVREVMRDATAK--V 228
OY 221 SAFFSAETTELSCGVVALATDPNILLSGVLPSCDLARRIGLRVDG--RP----- 271
DB 229 PHFIISETPRFIRAVVAAALADTDRSRMNGQSLSSGLAQQVGTDLDSGRPDAMRYPE 288
OY 272 VQD 274
DB 289 VQD 291

RESULT 12
O9EWT2 PRELIMINARY: PRT: 308 AA.
AC 09EWT2;
DR 01-MAR-2001 (TREMBLrel. 16, Created)
DR 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
GN PUTATIVE OXIDOREDUCTASE.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_Taxid=1902;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Parkhill J., Barrett B.G., Rajandream M.A.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Klenesch H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL449216; CAC14941.1;
SQ SEQUENCE 308 AA; 33485 MW; D7BD8D12D812832B CRC64;

Query Match
Best Local Similarity 20.7%; Score 335.5; DB 2; Length 308;
Matches 94; Conservative 55; Mismatches 110; Indels 43; Gaps 11;

OY 4 PMNGQCVVVGASRGIGGIALQCKAGATYITGRHL-----DLRVAAQEAQS 53
DB 8 PLAGRVAVAGATNRGAGGIVAGATVYVGRSTVRSEYDRRETIETADLTVE 67
OY 54 LGGCVVAVCDSSQSESERTLFEOVDREOGRDLVNNAAVAGOTILNTNKA-FWETPA 113
DB 68 AGGCGVAVPDDHDPDQVAVVNDRIASE-QARLDILVNDING--ETFE-----WDSF- 118
OY 114 SMQDINNNGLR-----GHYCSYVGARLMPVAGGGLIVVSSPSGL-----YMFNV 162
DB 119 -VVEHNDKGLRLRLAVETHAITSHALPDLRRPGVIVEVDGDAVNRDRSFF 177
OY 163 YGVGKACDKLADCAHELRHGVSCVSLMPGIYOTELKEHMAKEEV-LQDPVLKQFKS 221
DB 178 YDLAKTSVLEMAFSLGIEVPRGATAVALTPGMIRSEIMDHFGEVRENNRDLDR--VP 235
OY 222 SAFFSAETTELSCGVVALATDPNILLSGVLPSCDLARRIGLRVDG--RP-----V 272
DB 236 HFAISETPRVGVAVRRLAADPGVARNGRSFGSLAREYGTDLDSRDPAMRYLEV 295
OY 273 QD 274
DB 296 QD 297

RESULT 13
O9I2R7 PRELIMINARY: PRT: 255 AA.
AC 09I2R7;
DR 01-MAR-2001 (TREMBLrel. 16, Created)
DR 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
GN PROBABLE SHORT-CHAIN DEHYDROGENASE.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_Taxid=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PAOI;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Ervin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Bittman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
RA Garner R.L., Goltzy L., Tolentino E., Westbrock-Madman S., Yuan Y.,
RA Brody L.T., Coulter S.N., Folger K.R., Kas A., Lazdig K., Lam R.M.,

```



```
Db      9  L HGRRALITGASTGIGKRVALAYEAGAOYALIAARHIDALEKLADEIGTSGKVPVCCD 68
QY      65  SSOSEVYRTLEEQYDRQOGRLDVYVNNATAGVOTIINTRNKAFWETPASMDDINNVL 124
Db      69  VSOHQOVTSMIDOVTAELGIDIVACN--AGITTV-----TPILDWPLEEFORLQNTNV 120
QY      125 RGHYFCSVYGARLMVPAGOGIIVYISSPSGLQYMFNP-----YGVKAACDKLADCAH 179
Db      121 TGVFLTAQAATAKAMVKOGG-GVIINTASMSGHIINVPQOVSHYCASKAAVHILTKAMAV 179
QY      180 ELRRHGVCVSLMPGIYQFELLKEH 204
Db      180 ELAPHKIRVNSVSPGYILTELVEPY 204
```

Search completed: July 31, 2002, 15:12:39
Job time: 223 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:05:36 ; Search time 17.53 Seconds

(without alignment)
1715.685 Million cell updates/sec

Title: US-10-006-163-1

Sequence: 1623
1 MAAPMGQCVTVGASRGIG.....YLPSPFLRVPRKMTALYTSKF 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: PIR.71:*
- 2: PIR2:*
- 3: PIR3:*
- 4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	617.5	38.0	339	AE2212	hypothetical prote
2	615	37.9	323	T32125	hypothetical prote
3	583	35.9	325	T29604	hypothetical prote
4	347.5	21.4	326	T34594	probable oxidoredu
5	343.5	21.2	302	B83701	hypothetical prote
6	310	19.1	255	D83415	probable short-cha
7	286.5	17.7	248	AC1657	glucose 1-dehydrog
8	284.5	17.5	248	AH1285	glucose 1-dehydrog
9	271	16.7	255	D70635	hypothetical prote
10	262	16.1	250	B69802	glucose 1-dehydrog
11	254	15.7	246	A69621	3-oxoacyl-(acyl)-ca
12	247	15.2	249	B83767	glucose 1-dehydrog
13	247	15.2	266	AA2468	glucose 1-dehydrog
14	246	15.2	251	AA2042	7alpha-hydroxyster
15	243	15.0	266	F83127	3-oxoacyl-(acyl)-ca
16	241	14.8	249	F97338	probable short-cha
17	234.5	14.4	247	AG1672	3-ketoacyl-acyl ca
18	231.5	14.3	247	AG1300	3-ketoacyl-acyl ca
19	229	14.1	271	S34678	short-chain alcoh
20	226	13.9	260	B48674	trypsinone reducta
21	224.5	13.8	258	A96013	probable oxidoredu
22	224.5	13.8	278	T10877	Y41A protein - Rhl
23	224	13.8	224	H70447	3-oxoacyl-(acyl)-ca
24	222.5	13.7	256	A57149	glucuronate 5-dehydr
25	220	13.6	256	F97530	probable gluconat
26	220	13.6	256	AG2749	glucuronate dehydr
27	219	13.5	261	H98316	hypothetical prote
28	219	13.5	261	AD2966	dehydrogenase Atus
29	217.5	13.4	246	H72219	3-oxoacyl-(acyl) ca

30	216.5	13.3	278	E87642	short chain dehydr
31	216	13.3	246	C83961	3-oxoacyl-(acyl)-ca
32	215	13.2	247	S77280	3-oxoacyl-(acyl)-ca
33	214.5	13.2	246	B89896	3-oxoacyl-(acyl)-ca
34	214.5	13.2	248	B97223	probable 3-ketoacy
35	214	13.2	274	G83284	probable short-cha
36	214	13.2	320	S22450	3-oxoacyl-(acyl)-ca
37	213.5	13.2	275	I40211	probable sterol de
38	213	13.1	256	AD3235	2-deoxy-D-gluconat
39	212.5	13.1	281	F69400	short-chain alcoh
40	211.5	13.0	263	T07698	oxidoreductase, sh
41	211	13.0	258	H75616	3-oxoacyl-(acyl)-ca
42	210.5	13.0	247	B83880	protein F14N23.19
43	210	12.9	242	C86237	3-oxoacyl-(acyl)-ca
44	210	12.9	254	AC3256	3-oxoacyl-(acyl)-ca
45	209.5	12.9	251	AC0103	2-deoxy-D-gluconat

ALIGNMENTS

AE2212	hypothetical protein alr3252 [Imported] - Anabaena sp. (strain PCC 7120)
C:Species:	Anabaena sp.
A:Note:	Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date:	14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession:	AE2212
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irlig, N.; Nakazaki, N.; Shimpoo, S.; Sugimoto, M.; Takasawa, M.; Yamada, M.; Yasuda, M.; Tabata, DNA Res. 8, 203-213, 2001	
A:Title:	Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number:	AB1807; MUID:21595285; PMID:11759840
A:Accession:	AE2212
A:Status:	preliminary
A:Molecule type:	DNA
A:Residues:	1-339 <KUR>
A:Cross-references:	GB:BA000019; PIRN:BA074951.1; PID:g17132347; GSPDB:GN00179
A:Experimental source:	strain PCC 7120
C:Genetics:	
A:Gene:	alr3252
Query Match	38.0%; Score 617.5; DB 2; Length 339;
Best Local Similarity	43.0%; Pred. No. 1e-42;
Matches 142; Conservative 53; Mismatches 94; Indels 41; Gaps 12;	
QY 5 MNGQVCVTVGASRGIGRGIALGKAGATVYIRGRHLDT-----LRVVAQE 50	
DB 16 LEKVALVVGATGKLGKIALIGEGALVYITGRRLDNNSTSGNDVSSGLNETKIAVEE 75	
QY 51 AASLGQCVTVVCDSSQSESEVRTLFEVDREDOGRLDVLYNNAYAGVQTLNTRNAKAFE 110	
DB 76 A---GGICIFPVGDHNSNDQVRLFERIQENGQGLDLVNNAYAGVQALFTNAQGRPEWE 132	
QY 111 TRPSMMDIDINNVGLRHGFCFSYVGARLMPAAGGLVYISSPSSLOYMNPVPGVCAKAC 170	
DB 133 NEPSLWDACNNWGLRSHYTLASVYAAMMSKRQGLICTTSSWGMAYLFGAAYGAKAC 192	
QY 171 DKLAACAHLELRHGVSCVSLMPGIVOTELRKEMAKEEVLQDPVLRKOPK-SAFSSA--- 226	
DB 193 DRLAAMAYELQYNNASLSIMPQIVGTLEF-SRLASE--MSDNTDNGKNSALIRERYNW 249	
QY 227 ETTLSGKCVVALATDPNIIISGKVLPSCDLARRYGLRDVDR-PVQ-----DYLSL 278	
DB 250 EFTPLLTGRVIAKLAETVNIIFRQVVAELAKQYSLVDQEGNLPVLSRLSRLFLPLAL 309	
QY 279 SSVLSHVSGLGWLASTLPSTLRVPMKIAL 308	
DB 310 PTLRKH----SWL----IPD-IKVP-WSILL 330	

RESULT 2

T32125

hypothetical protein F59E11.2 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

R:Bradshaw, H.
 Submitted to the EMBL Data Library, July 1997

A:Description: The sequence of *C. elegans* cosmid F59E11.
 A:Reference number: 221124

A:Accession: T32125

A:Molecule type: DNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Cross-references: EMBL:AF016685; PIDN:AA66216.1; GSPDB:GN00023; CESP:F59E11.2

A:Gene: CESP:F59E11.2

A:Map position: 5

A:introns: 20/1; 60/1; 108/3; 135/2; 166/2; 192/3; 227/3; 280/1

Query Match

Best Local Similarity 37.9%; Score 615; DB 2; Length 323;
 Matches 138; Conservative 58; Mismatches 112; Indels 20; Gaps 5;

1 MAPMNGCVVVTGASRGIGRGIALQICRAGATVYITGRH-----LDTLRVVAOE 50
 1 MGVITDDOVALYVGASRGIGRGIALQIGEGANVYITGRPELSDNFRIGLPSLDVVAKE 60
 51 AOSLGQCVPVVCDSSQSESEVRLPEQVDRQOGLDLVNNAYAGVOTILNTNKAFAWE 110
 61 ITRSGKIGALIVYDSHNMTEVEFLFEKEDDEKGLDILVNNVYSLGKATMIGKTFED 120
 111 TPASMDINNNGVLRGHYCSYVYGARLWPAQOGLIVYSSGSIQYFNVPYGVGAAC 170
 121 ODPSTMDINNGVLRGHYCSYVYGARLWPAQOGLIVYSSGSIQYFNVPYGVGAAC 180
 171 DKLADCAHELRHRCVSCVSLMPGIVOTELKEHM---AKEEVLDDPLKQFSSAE 227
 181 ARMSDVAEELNPNVNCVVTILPGVKTETANRPIIDAYKMEPELEE---FIRGE 236
 228 TTLSGKCVVALATDPNLTSLSGVLPSCDLARRYGLRDVDGRPV---QDYLSSTVSHV 285
 237 STEYTGKALRLANDPKLTKSGTLFTEDLAOKYDFSDKHGAGMEPQNRISITILG-T 295
 286 SGLGWLASYLPSFLRVPKRIATLYTSKF 313
 296 MKEEVAKYIPQIKLPKWIQSVNRF 323

RESULT 3

T32604

hypothetical protein ZK816.5 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

R:Nhan, M.; Le, T.
 submitted to the EMBL Data Library, November 1995

A:Description: The sequence of *C. elegans* cosmid ZK816.
 A:Reference number: 220649

A:Accession: T32604

A:Molecule type: DNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Cross-references: EMBL:U41018; PIDN:AA082327.1; CESP:ZK816.5

A:Gene: CESP:ZK816.5

A:introns: 24/1; 111/3; 170/2; 228/3; 280/1

Query Match

Best Local Similarity 35.9%; Score 583; DB 2; Length 325;
 Matches 130; Conservative 55; Mismatches 111; Indels 36; Gaps 6;

5 MNGOVCVVVTGASRGIGR-----GIALQICRAGATVYITGR-----HDTLRVVAO 49

3 LKSLAIYVVGASRGIGRGLMKVLAQAEAGTLYTGAPSKTSLSELYPTLEGTA 62

50 EAOSLGQCVPVVCDSSQSESEVRLPEQVDRQOGLDLVNNAYAGVOTILNTNKAFAW 109

63 ECRRGGICVRYVDHSMNDEVEKFEDEVASETDQDLILVNNAYAVTKGSGDTRRF 122

110 ETPASMDINNNGVLRGHYCSYVYGARLWPAQOGLIVYSSGSIQYFNVPYGVGA 168

123 ENDEPITDDINNVGLRNQYCSYVYGARLWPAQOGLIVYSSGSIQYFNVPYGVGA 182

169 ACDRIADCAHELRHRCVSCVSLMPGIVOTELKEHM-----AKEEVLDDPLKQF 220

183 ALDRMSDVAEELNPNVNCVVTILPGVKTETANRPIIDAYKMEPELEE---FIRGE 236

221 SAFSSATTELSGKCVVALATDPNLTSLSGVLPSCDLARRYGLRDVDGRPV---QDYLSSTVSHV 285

234 --FLNGSTELSGKCVVALATDPNLTSLSGVLPSCDLARRYGLRDVDGRPV---QDYLSSTVSHV 285

281 VLSHVSGLGWLASYLPSFLRVPKRIATLYTSKF 312

292 LLS-LAGYHSMAGMCEPWNLPQMAITLWQNK 322

RESULT 4

T34594

Probable oxidoreductase - *Streptomyces coelicolor*

C:Species: *Streptomyces coelicolor*
 C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999

R:Munphy, L.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, January 1998

A:Accession: T34594

A:Molecule type: DNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Cross-references: EMBL:AL021529; PIDN:CA016459.1; GSPDB:GN00070; SCODEB:SC10A5.27C

A:Gene: SCODEB:SC10A5.27C

Query Match 21.4%; Score 347.5; DB 2; Length 326;
 Best Local Similarity 33.6%; Pred. No. 1e-20;
 Matches 99; Conservative 55; Mismatches 102; Indels 39; Gaps 11;

4 PMNGOVCVVVTGASRGIGRGIALQICRAGATVYIT-----GRHDTLRVVAOEAQ 52

22 PLAGRIALVAGATRGAGRAOAVELGRAGATVYITGTRARASEVGRTEETIEYELV 81

53 SLGQCVPVVCDSSQSESEVRLPEQVDRQOGLDLVNNAYAGVOTILNTNKAFAW 110

82 AAGGCIATVPTTHLDEAQRALVERIDREYE--RLDILVDINGG--EHLATSVFGKSW 139

111 TPASMDINNNGVLRGHYCSYVYGARLWPAQOGLIVYSSP-----GSLQYWF 159

140 TPLADGLRIELGASH---VITALLP-----LLIRSDADLHVETDGTASHRRRYE 191

160 NVPYGVGAACDKLAADCAHELRHRCVSCVSLMPGIVOTELKEHM--AKEEVLDDPLKQ 218

192 NIYDLAKNAPRIALFAGLAELAEYGAIVASPEFLSEQMLSHFGVSEEMRDAIOE 251

219 FKSAFSAETTELSGKCVVALATDPN--ILSLSGVLPSCDLARRYGLRDVDGRV--RP 271

252 --PTFAIESPHYLARTVAALADPDRAKRWNGSTSGELARAYGVYDVGSRP 304

RESULT 5

B83701

hypothetical protein BH0410 [imported] - *Bacillus halodurans* (strain C-125)

C:Keywords: NAD, oxidoreductase
F:6-108/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 15.2%; Score 247; DB 1; Length 266;
Best Local Similarity 29.7%; Pred. No. 1.2e-12;
Matches 78; Conservative 42; Mismatches 107; Indels 36; Gaps 7;

QY 5 MNGOVCVYTGASRGIGRGIALQCKAGATVYITGRHLDLTVVAQEAOSLGQCVPVCD 64
DB 3 LKDKVILVASTRGIGLALQAQAKGAKVYMGARLERAKADEMNAAGGNVXYVND 62
QY 65 SSGSESEVTLFEQVDREOQGRDLVNNAYAGVCT---ILNTRKAWMPERPAWMDIN 120
DB 63 ATKEETVYTMIEEI-IEQGRIDLVLVNNFGSSNPKKDGIANITDEPFIKTV----- 113
QY 121 NVGRGHVFCVYGARLNVAPAGGLIYVSSPSGL-QYMFVVPYGVGAACDKLAADCAH 179
DB 114 NINIKSVFIASQTVAKYMAENGSGSIINISSVGGILIPDISQIAGVTSAAINLYLKLIAV 173
QY 180 ELRRHGVSCVSLMPGIYOTELKREMAKEEVLDPVAKORSAFSSAETTELSC----- 233
DB 174 HEARNIRCNANVLEGMAT-----DAVODNLTDPRNFELKHTPIORMGLPEEIA 223
QY 234 KCVVALATDP-----NLSLSG 250
DB 224 AAVYFASDDAAVYTGQILTVSG 246

RESULT 14

AH2042
3-oxoacyl-[acyl-carrier protein] reductase [Imported] - Anabaena sp. (strain PCC 7120)
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
R:Accession: AH2042
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kurlitz, T.; Sasamoto, S.; Watanabe, A.; Itiguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete genomic sequence of the filamentous nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2042
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-251 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA873593.1; PID:917130984; GSPDB:GN00179
C:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: fadG
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 15.2%; Score 246; DB 2; Length 251;
Best Local Similarity 31.0%; Pred. No. 1.3e-12;
Matches 85; Conservative 39; Mismatches 106; Indels 44; Gaps 11;

QY 5 MNGOVCVYTGASRGIGRGIALQCKAGATVYITGRHLDLTVVAQEAOSLGQCVPV 62
DB 8 LRGVAAVYTGASRGIGRGIALQCKAGATVYITGRHLDLTVVAQEAOSLGQCVPV 62
QY 63 CDSOSESEVTLFEQVDREOQGRDLVNNAYAGVCT---ILNTRKAWMPERPAWMDIN 116
DB 67 ADVSQ-----VEQVDNLNGALDKRRIIDLVNN--AGI-----TRDTLLLRMRPEW 112
QY 117 DDINNGLRGHYFCVYGARLNVAPAGGLIYVSS-PSGLQYMFVVPYGVGAACDKLA 175
DB 113 QAVIDLTLTGVLCTRAVSLMKKORSGRIINTSVAGOMGNPGQANYSAKAGVIGFTR 172
QY 176 DCACHELRHGVSCVSLMPGIYOTELKREMAKEEVLDPVAKORSAFSSAETTELSC 235
DB 173 TVAKELASRGITVNAVAPGIATD-MTSLNLSKSGILQYIPIGRY-----GOPEETAG-M 224
QY 236 VVALATDPNILLSKGVLPSCDLARRYGLRDVDG 269

DB 225 VRFIADPAAAYITGOVF-----NVDG 246

RESULT 15

Probable short-chain dehydrogenase PA4148 [Imported] - Pseudomonas aeruginosa (strain F83127)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83127
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laddig, K.; Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337.
A:Accession: F83127
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-266 <STO>
A:Cross-references: GB:AE004831; GB:AE004091; NID:9950347; PIDN:AG07535.1; GSPDB:G
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4148
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 15.0%; Score 243; DB 2; Length 266;
Best Local Similarity 32.5%; Pred. No. 2.5e-12;
Matches 87; Conservative 30; Mismatches 99; Indels 52; Gaps 8;

QY 5 MNGOVCVYTGASRGIGRGIALQCKAGATVYITGRHLDLTVVAQEAOSLGQCVPVCD 64
DB 11 LSRVALVLTGAGGIGRGIALALRAGADVAVDLPQVAEETRAAATRSLSRLAIGVD 70
QY 65 SSGSESEVTLFEQVDREOQGRDLVNNAYAGVCT---ILNTRKAWMPERPAWMDINNVGL 124
DB 71 VSDGDSVAMVERVARE-FGRLDVAVNNA-----GVISIRKVA--ELSLADMVRVNVNA 122
QY 125 RGHYFCVYGARLNVAPAGGLIYVSSPSGLQYMFVVPYGVGAACDKLAADC----- 177
DB 123 RGVFLCCQAEELPLMQAQRNGRIYVNLSSI-----AGRVGLPDLAHTCASFAYI 170
QY 178 -----AHELRHGVSCVSLMPGIYOTELK-----EHMAKE-EVLDPVTK 217
DB 171 GSNALAKEVARDGVTVNALCPGIYGTGMWRGEGDLSGRMROAGESEAGSWERHQAISLP 230
QY 218 QFSAFSSAETTELSCGVVALATDPNI 245
DB 231 Q-----GEAQTVEDMGQLVYLACAPHV 253

Search completed: July 31, 2002, 15:09:16
Job time: 220 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:09:21 ; Search time 13.44 Seconds

(without alignments)
901.728 Million cell updates/sec

Title: US-10-006-163-1

Perfect score: 1623

Sequence: 1 MAPPMNGVCVVTGASRGIG.....YLPSELRVPRKIALYTSKF 313

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	254	15.7	246	1 FABG_BACSU
2	229	14.1	271	1 SDR1_PICAB
3	226	13.9	260	1 TRN2_DATST
4	224	13.8	278	1 YALA_RHISN
5	224	13.8	248	1 FABG_RHISN
6	222	13.7	256	1 GNO_GLUOX
7	217	13.4	246	1 FABG_THENX
8	215	13.2	247	1 TRN2_HYONI
9	214	13.2	320	1 FABG_CUPLA
10	214	13.2	275	1 YCP1_BRAJA
11	213	13.2	275	1 YCP1_BRAJA
12	209	12.9	256	1 Y019_THEMA
13	206	12.7	319	1 FABG_ARATH
14	204	12.6	262	1 YXBG_BACSU
15	203	12.5	255	1 HDHA_ECOLI
16	200	12.3	238	1 YORD_BACSU
17	193	11.9	313	1 YXEK_GAEEL
18	192	11.8	592	1 EPHD_MYCRU
19	191	11.8	251	1 Y325_THEMA
20	190	11.7	261	1 ACT3_STRCO
21	190	11.7	267	1 HDHA_CLOSO
22	189	11.7	256	1 BDDC_KLEPN
23	188	11.6	247	1 FABG_PSEAE
24	186	11.5	336	1 TSD_MAIZE
25	185	11.4	253	1 KDDP_ECOLI
26	185	11.4	264	1 UCDA_SALRY
27	185	11.4	244	1 FABG_VIBHA
28	185	11.4	258	1 BDHA_RHIME
29	185	11.4	272	1 DRL1_STRVN
30	184	11.4	253	1 KDDP_ERMCH
31	183	11.3	242	1 FABG_ACTAC
32	183	11.3	248	1 YXEK_RHISN
33	183	11.3	261	1 DH8B_HUMAN

34	182.5	11.2	263	1 UCDA_ECOLI	P37440 escherichia
35	181	11.2	247	1 YD50_MYCCTU	Q11020 mycobacteri
36	181	11.2	261	1 BEND_ACICA	P07772 a cis-1,2-d
37	180.5	11.1	254	1 IDNO_ECOLI	P39345 escherichia
38	179.5	11.1	242	1 FABG_HAEIN	P43713 haemophilus
39	179	11.0	260	1 DH8B_MOUSE	P50171 mus musculu
40	176	10.8	241	1 BDDC_KLETE	Q04520 klebsiella
41	176	10.8	244	1 FABG_VIBCH	Q9K9H7 vibrio chol
42	176	10.8	254	1 KDDP_BACSU	P50842 bacillus su
43	175.5	10.8	261	1 DHG_BACME	P40288 bacillus me
44	175	10.8	244	1 FABG_ECOLI	P25715 escherichia
45	173.5	10.7	262	1 DH8B_BACME	P07999 bacillus me

ALIGNMENTS

RESULT 1

FABG_BACSU STANDARD: PRT; 246 AA.

ID FABG_BACSU

AC P51831; O31733.

DT 01-OCT-1996 (Rel. 34, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) (3-ketoacyl-acyl carrier protein reductase).

GN FABG.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

CC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_TaxId=1423;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=96326321; PubMed=8759840;

RA Morbidoni H.R., de Mendoza D., Cronan J.E. Jr.:

RT "Bacillus subtilis acyl carrier protein is encoded in a cluster of lipid biosynthesis genes."

RT J. Bacteriol. 178:4794-4800(1996).

RN (2)

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98195738; PubMed=9534248;

RA Foulger D., Errington J.:

RT "A 28 kbp segment from the spoVA region of the Bacillus subtilis 168 genome."

RT Microbiology 144:801-805(1998).

RN (4)

RP SEQUENCE OF 230-246 FROM N.A.

RC STRAIN=168;

RX Oguro A., Takeshita H., Takamatsu H., Yamane K.:

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] + NAD(+) -> 3-oxoacyl-[acyl-carrier protein] + NADPH.

CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS PATHWAY.

CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to license@isb-sdb.ch).

CC -----

DR EMBL, U59433; AAC4307.1; -.

DR EMBL: 299112; CAB13464.1; -
 DR EMBL: Y13937; CA74250.1; -
 DR EMBL: D64116; BA10974.1; -
 DR HSSP: P50162; IABF.
 DR Subtilisin; B61155; fabc.
 DR Interpro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short.1.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 DR Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome;
 NP_BIND 25
 FT ACT SITE 154 154 NADP (BY SIMILARITY).
 FT CONFLICT 23 23 D -> A (IN REF. 1).
 SQ SEQUENCE 246 AA; 26282 MW; C6A391167D3237DC CRC64;

Query Match 15.7%; Score 254; DB 1; Length 246;
 Best Local Similarity 29.8%; Pred. No. 6.7e-14;
 Matches 77; Conservative 46; Mismatches 105; Indels 30; Gaps 11;

QY 5 MNGQCVYTGASRGIGRIALQCKAGATVY--TGRHLDTLRYVAQASLGQCVPV 62
 Db 2 LNDKTAIVTGAASRGISIALDLAKSANYVYVSGNEAKANEYV-DEIKSWGRALAVK 60
 QY 63 CDSOSEVETLFEQVDREQGRDLVYNNNAVAGVQITLNRKAFETPASMDDINNV 122
 Db 61 ADVSPPEVQMKIET-LVSTFIDILVNN--AGI-----TRDLIMKEDDWDVYNI 112
 QY 123 GLRGHYFCYSYGARLPAPAGGLIVYISS-----PSGLQMFNVFPGVGAACDKIAA 175
 Db 113 NLKGVNCTKAVTRQMKMGRGRIINVSIVGSGNPGQANYV-----AAKAGVIGLTK 166
 QY 176 DCAEHLRHGVCSCVSLMPGIVQTELKEHMAKEVYLDPVLKQKFAESSAETTELSC 235
 Db 167 SSAKELASRNTVNAIAGFISTD--VDEMILKQIPLA-RFGEPDVS- 220
 QY 236 VVALATDPNIIISLGSKVL 253
 Db 221 VTFLEASGARVYMTGTL 238

RESULT 2
 ID SDR1_PICAB STANDARD; PRT; 271 AA.
 AC 008633;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Short-chain type dehydrogenase/reductase (EC 1.-.-.-).
 OS Picea abies (Norway spruce) (Picea excelsa).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
 NC NCBL_TaxID=3329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94120027; PubMed=8290650;
 RA Bauer S., Galliano H., Pfeiffer F., Messner B., Sandermann H.,
 RA Ernst D.;
 RT "Isolation and characterization of a cDNA clone encoding a novel
 RT short-chain alcohol dehydrogenase from Norway spruce (Picea abies L.
 RT Karst)."
 RL Plant Physiol. 103:1479-1480(1993).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: X74115; CAAS2213.1; -
 DR HSSP: Q12634; IYBV.
 DR Interpro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short.1.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 DR Oxidoreductase.
 NP_BIND 25
 FT ACT SITE 179 179 NAD OR NADP (BY SIMILARITY).
 FT CONFLICT 179 179 BY SIMILARITY.
 SQ SEQUENCE 271 AA; 28724 MW; 743EALCB2F85FE3 CRC64;

Query Match 14.1%; Score 229; DB 1; Length 271;
 Best Local Similarity 30.8%; Pred. No. 8.6e-12;
 Matches 77; Conservative 42; Mismatches 107; Indels 24; Gaps 10;

QY 4 PMNGQCVYTGASRGIGRIALQCKAGATV--YITGRHL--DTLRYVAQASLGQCV 58
 Db 18 PLGGVVAIVTGAASRGIGRIALNMAEKGVYIHYSSNOHAAEEVASTINNKSPSSGDCV 77
 QY 59 VPVVC--DSSOSEVETLFEQVDREQGRDLVYNNNAVAGVQITLNRKAFETPASM 116
 Db 78 RAIVCKADVAEPQVAQLFDTRAE-HAFGPLHIVNNA-----GVYDSKPYTLAOTSDEW 131
 QY 117 DDINNVGLRGHYFCYSYGARLPAPAGGLIVYISSPSGLQMFNVF-----YGVGRACDK 172
 Db 132 DRIVQNVCKGAFGLCSREAKAVYRGGGRRIINIS--SLVAN-PIPRGATITASAAYEM 188
 QY 173 LAACAEHLRHGVCSCVSLMPGIVQTELKEHMAKEVYLDPVLKQKFAESSAETTEL 232
 Db 189 MTRILAQLRGTOITANCAVPATDMF--FAGKSEAVERGVK--SNPFERLQKVEDV 244
 QY 233 GKCVVALATD 242
 Db 245 APLVAFPLASD 254

RESULT 3
 ID TRN2_DATST STANDARD; PRT; 260 AA.
 AC P50163;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tropinone reductase-II (EC 1.1.1.236) (TR-II).
 GN TR2.
 OS Datura stramonium (Jimsonweed) (Common thornapple).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Datura.
 NC NCBL_TaxID=4076;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Root;
 RX MEDLINE=94022421; PubMed=8415746;
 RA Nakajima K., Hashimoto T., Yamada Y.;
 RT "Two tropinone reductases with different stereospecificities are
 RT short-chain dehydrogenases evolved from a common ancestor".
 RL Proc. Natl. Acad. Sci. U.S.A. 90:9591-9595(1993).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=98226735; PubMed=9560196;
 RA Hashimoto T., Oda J., Yamada Y.;
 RT "Crystal structures of two tropinone reductases: different reaction
 RT stereospecificities in the same protein fold."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:4876-4881(1998).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE=99316165; PubMed=10387002;
 RA Yamashita A., Kato H., Wakatsuki S., Tomizaki T., Nakatsu T.,
 RA Nakajima K., Hashimoto T., Yamada Y., Oda J.;
 RT "Structure of tropinone reductase-II complexed with NADP+ and

```

RT pseudotriple at 1.9 A resolution: implication for stereospecific
RL substrate binding and catalysis."
CC Biochemistry 38:7630-7637(1995).
CC -1- FUNCTION: CATALYZES THE STEREOSPECIFIC REDUCTION OF TROPINONE TO
CC PSEUDOTROPINE.
CC -1- CATALYTIC ACTIVITY: Pseudotriple + NADP(+) -> tropinone + NADPH.
CC -1- PATHWAY: BIOSYNTHETIC PATHWAY OF TROPANE ALKALOIDS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: L20474; AAA33282.1; -
CC PDB: 2AE1; 18-NOV-98.
CC PDB: 2AE2; 02-FEB-99.
CC InterPro: IPR002198; ADH_short.
CC Pfam: PF00106; adh_short; 1.
CC PRINTS: PR00080; SDRFAMILY.
CC PROSITE: PS00061; ADH_SHORT; 1.
CC Oxidoreductase; NADP; 3D-structure.
CC NP_BIND 13 37 NADP (BY SIMILARITY).
CC ACT_SITE 159 159 BY SIMILARITY.
CC SEQUENCE 260 AA; 28311 MW; 2DBF4963B2CA303 CRC64;

Query Match 13.9%; Score 226; DB 1; Length 260;
Best Local Similarity 28.7%; Pred. No. 1.5e-11;
Matches 70; Conservative 39; Mismatches 113; Indels 22; Gaps 6;

QY 5 MNGCVVTVGASRGIGRGIALQCKAGATVYITGRHLDLTVVAOAOAGSLGGCCVPVCD 64
DB 7 LEGCALVTGSGRGIGYVEELASGASVYTCRNOKELNDCLTMRKSGFVEVASVCD 66
QY 65 SSOSEVRLTEFQVDRDQGRDLVNNVAGVQTILNTRNKAFTWTPASMDINDINVL 124
DB 67 ISSREBRELMTVANHFHGLINTLVNN--AGI--VIYKEAKDYTEDYSL--INSIN 119
QY 125 RGHYFCVYGARLMPVAGGGLVIVISS-PGSLQYMENVYGVGKACDKLADCAHELR 183
DB 120 EAAVHLSTVLAHPFLKASERGNVVFISVSGALAVPEAVYGAATKGAMDLTICLAFEMAK 179
QY 184 HGVSCVSLMPGIVQTELLKEHNAKEEVLDPVLKOPKSAFSSAETTELSCVVALATDP 243
DB 180 DNIRNVGVPGVYATISLV-----EMTIODPECKE-----NLNKLIDRCALRMGEP 225
QY 244 NILS 247
DB 226 KEIA 229

RESULT 4
YALA_RHISN STANDARD; PRT; 278 AA.
ID YALA_RHISN
AC P55541;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative short-chain type dehydrogenase/reductase, YALA (EC 1.-.-.-).
GN YALA.
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym PNGR234a.
CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
CC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN (1)
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE-97305956; PubMed-9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes."
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: AE000082; AAB91754.1; -
CC HSSP: P50163; 2AE1.
CC InterPro: IPR002198; ADH_short.
CC Pfam: PF00106; adh_short; 1.
CC PRINTS: PR00080; SDRFAMILY.
CC PROSITE: PS00061; ADH_SHORT; 1.
CC Hypothetical protein; Oxidoreductase; Plasmid.
CC NP_BIND 10 34 NAD OR NADP (BY SIMILARITY).
CC ACT_SITE 159 159 BY SIMILARITY.
CC SEQUENCE 278 AA; 28743 MW; 1D0105625BE9DF2B CRC64;

Query Match 13.8%; Score 224.5; DB 1; Length 278;
Best Local Similarity 31.0%; Pred. No. 2.1e-11;
Matches 67; Conservative 34; Mismatches 98; Indels 17; Gaps 6;

QY 7 GOVCVTVGASRGIGRGIALQCKAGATVYITGRHLDLTVVAOAO--SLGGCCVPVCD 64
DB 6 GKAAVVTGAGAGIGKACALAIAREGGRVVA--DLGSAIACTAOIAEAGNALAMND 63
QY 65 SSOSEVRLTEFQVDRDQGRDLVNNVAGVQTILNTRNKAFTWTPASMDINDINVL 124
DB 64 IADAQVVAALFEFAER-HRGVDLVNNNSA--MLTPDRRLTDLDAVMDQTMATLV 119
QY 125 RGHYFCVYGARLMPVAGGGLVIVISS-PGSLQYMENVYGVGKACDKLADCAHELR 183
DB 120 RGTLLCCRQAIIPMIRAGGAIYVNMSSCGLSGDTAQTGSAVSKAMNLSLAVOYCH 179
QY 184 HGVSCVSLMPGIVQTELL-----KEHNAKEEVL 211
DB 180 AQIRCNVAVPGLIMTERLAKIDCKMQRHLSRHQL 215

RESULT 5
FABG_AQUAE STANDARD; PRT; 248 AA.
ID FABG_AQUAE
AC O67610;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) (3-ketoacyl-
DE acyl carrier protein reductase).
GN FABG OR AQO_1716.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE-98196666; PubMed-9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
CC Nature 392:353-358(1998).
CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +

```

```

CC      NADP(+) - 3-oxoacyl-[acyl-carrier protein] + NADPH.
CC      -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
CC      PATHWAY.
CC      -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC      (SDR) FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC      or send an email to license@sdb.ch).
CC      -----
DR      EMBL: AE000752; AAC07575.1; -.
DR      HSSP: P50163; 2AE1.
DR      InterPro: IPR002198; ADH_short.
DR      Pfam: PF00106; adh_short; 1.
DR      PRINTS: PR00080; SDRFAMILY.
DR      PROSITE: PS00061; ADH_SHORT; 1.
KW      Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT      NP_BIND 12 36 NADP (BY SIMILARITY).
FT      ACT_SITE 157 157 BY SIMILARITY.
SQ      SEQUENCE 248 AA; 26867 MW; 5CFD9EB9AD83F2C5 CRC64;

Query Match      13.8%; Score 224; DB 1; Length 248;
Best Local Similarity 30.8%; Pred. No. 2e-11;
Matches 72; Conservative 37; Mismatches 103; Indels 22; Gaps 7;

QY      1 MAPNMGVCYVTGASRGIALQKAGATVITGRHDLTVNAQF-AQSLGGCV 59
DB      1 MEIKIAGKSVLTGTSTRGIRAIKALASGVITITGSEKRAVAEELANKKGVAAH 60
QY      60 PIVCSOSESEVETLEQVDREOGRDLVNNNAVAGVOTILTRNKAFWETPASMWDI 119
DB      61 GVEANMLSEESINKAFEEIYNLVDS-IDIVNN-AGI-----PDKFLKMSLIDMEV 112
QY      120 NNVGIRGHVFCGYARLWPAAGGLIYVIS-PCSLQYMFNVPGVAKACDKLAADCA 178
DB      113 LKVNLTGFLVYQNSLRKIKORMGRIVNISVYGFQNVGVNSTKALIGFTKSLA 172
QY      179 HELRRHGVSCVSLMPGIQVOTELKEHMAKEEVLDPVLKORK-----SAFSAAE 227
DB      173 KELARNVLVNNAVAPGFIETDK-----TAVLSEIKOKTKEQIPLGRFSPE 219

RESULT 6
GNO_GLUOX      STANDARD; PRT; 256 AA.
ID      GNO_GLUOX
AC      P50199;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Glucanate 5-dehydrogenase (EC 1.1.1.69) (5-keto-D-gluconate 5-
DE      reductase).
DE      GNO.
OS      Glucanobacter oxydans (Glucanobacter suboxydans).
OC      Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC      Glucanobacter.
OX      NCBI_TaxID=442;
RN      [1]
RP      SEQUENCE FROM N.A., SEQUENCE OF 3-19, AND CHARACTERIZATION.
RX      STRAIN=DSM 3503;
RC      MEDLINE=95270578; PubMed=7751271;
RA      Klaseen R., Bringer-Meyer S., Sahm H.;
RT      "Biochemical characterization and sequence analysis of the
RT      glucanobacter 5-oxido-reductase gene from Glucanobacter oxydans."
RL      J. Bacteriol. 177:2637-2643(1995)
CC      -1- FUNCTION: INVOLVED IN THE NONPHOSPHORYLATED, KETOGENIC OXIDATION
CC      OF GLUCOSE AND OXIDIZES GLUCONATE TO 5-KETOGLUCONATE. DEPENDENT ON
CC      NADP. ALMOST INACTIVE WITH NAD.
CC      -1- CATALYTIC ACTIVITY: D-gluconate + NAD(P)(+) -> 5-dehydro-D-

```

```

CC      gluconate + NAD(P)H.
CC      -1- SUBUNIT: HOMODIMER.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC      (SDR) FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC      or send an email to license@sdb.ch).
CC      -----
DR      EMBL: X80019; CA56322.1; -.
DR      HSSP: P47227; 1BDB.
DR      InterPro: IPR002198; ADH_short.
DR      Pfam: PF00106; adh_short; 1.
DR      PRINTS: PR00080; SDRFAMILY.
DR      PROSITE: PS00061; ADH_SHORT; 1.
KW      Oxidoreductase; NADP.
FT      NP_BIND 15 39 NADP (BY SIMILARITY).
FT      ACT_SITE 160 160 BY SIMILARITY.
SQ      SEQUENCE 256 AA; 27256 MW; 38B03C039C0A07A CRC64;

Query Match      13.7%; Score 222.5; DB 1; Length 256;
Best Local Similarity 31.1%; Pred. No. 2.8e-11;
Matches 66; Conservative 32; Mismatches 93; Indels 21; Gaps 5;

QY      5 MNGVCYVTGASRGIRGIALQKAGATVITGRHDLTVNAQF-AQSLGGCV 64
DB      9 LSGARALVTGASRGIGLGLANGIRGAEVYLVNGNASLSAQSGFAEGLKASTAVFD 68
QY      65 SSOSEVETLEQVDREOGRDLVNNNAVAGVOTILTRNKAFWETPASMWDI 124
DB      69 VTDDDAVIDGVAAIERD-MGPIDILINN-AGIQ-----RAPLEEFGRKMDMLSTNV 120
QY      125 RGHVFCVYGARLWPAAGGLIYVIS-PCSLQYMFNVPGVAKACDKLAADCA 177
DB      121 NAVFVGQAVARHHPKRGATVINCYSVOSELARGI-----APYATKGAVNLTGKM 174
QY      178 AHELRRHGVSCVSLMPGIQVOTELKEHMAKEE 209
DB      175 ATDWGRHGLQINGLAPGYFATEMTERLVADDE 206

RESULT 7
FABG_THEMA      STANDARD; PRT; 246 AA.
ID      FABG_THEMA
AC      Q9X248;
DT      30-MAY-2000 (Rel. 39, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) (3-ketoacyl-
DE      acyl carrier protein reductase).
DE      FABG OR TM1724.
GN      Thermotoga maritima.
OS      Thermotoga maritima.
OC      Bacteria; Thermotogales; Thermotoga.
OX      NCBI_TaxID=2336;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=MSB8 / DSM 3109;
RC      MEDLINE=99287316; PubMed=10360571;
RA      Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA      Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA      McDonald L., Utterback T.R., Malek J.A., Liner K.D., Garrett M.M.,
RA      Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA      Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA      Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT      "Evidence for lateral gene transfer between Archaea and Bacteria from
RT      genome sequence of Thermotoga maritima."
RL      Nature 399:323-329(1999).

```

```

CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
CC NADP(+) -> 3-oxoacyl-[acyl-carrier protein] + NADH.
CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
CC PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AE001811; AD36790.1; -.
CC DR HSSP: P50162; 1AE1.
CC DR TIGR: TM1724; -.
CC DR InterPro: IPR002198; ADH_short.
CC DR Pfam: PF00106; adh_short.1.
CC DR PRINTS: PR00080; SDRFAMILY.
CC KM Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
CC FT NP_BIND 10 34 NADP (By SIMILARITY).
CC FT ACT_SITE 154 154 BY SIMILARITY.
CC SQ SEQUENCE 246 AA; 26401 MW; 8C08904D28099142 CRC64;

Query Match 13.4%; Score 217.5; DB 1; Length 246;
Best Local Similarity 27.6%; Pred. No. 6.8e-11;
Matches 72; Conservative 51; Mismatches 101; Indels 37; Gaps 10;

QY 5 MNGGVCVVTGASRGIGRIALQLCAGATVY--ITGRHDLTRVAADEAGSLGGCCVPPV 61
DB 3 LGGVLLITGTAASIGATLTLLFADEGATVAGDSKENDLSL--VKEAEGLPKGVPPY 59
QY 62 VCDSSSEVETLFEQYDREOQGRDLVYNNAYAGVOTILTRNKAFETPASMDDINN 121
DB 60 VLNTDDQIEVEYEVK-VQKRGRIADVNN--AGI-----TRDLILRKKEEDMDAVIN 111
QY 122 VGLRGHFCVSYGARLWVPAGOGILIVISS-----PGLQYMNVPYGVGAACDKLA 174
DB 112 VNLKGVNVTQMVVPMKORNGSIVNVSIVGIYNGPQGN-----YAAKAGVIGMT 165
QY 175 ADCHELRRHGVSCVSLMPGVOTELKE--HMAKEVLODPVLKQFSAFSATTELS 232
DB 166 KTWAKELAGRIRVNAAPGFIETPMTEKLEPKARETALSRIPLRGFKPREVAGV--- 221
QY 233 GKCVVALATDPNILSLSGKVL 253
DB 222 ---ILFLASDESSY-VTGGV 238

RESULT 8
FAGL_SYNY3
ID FAGL_SYNY3 STANDARD; PRT: 247 AA.
AC P73574;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-oxoacyl-[acyl-carrier protein] reductase 1 (EC 1.1.1.100) (3-
DE ketoacyl-acyl carrier protein reductase 1).
GN TABG1 OR SLR0886.
OS Synchocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_Taxid=1148;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97061201; PubMed=8905231;
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugijura M., Sasamoto S., Kilmura T.,
RA Hosouchi T., Matsuo A., Muraki A., Nakazaki N., Naro K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tsubata S.;

```

```

RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
CC NADP(+) -> 3-oxoacyl-[acyl-carrier protein] + NADH.
CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
CC PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: D90907; BA117614.1; -.
CC DR HSSP: P50163; 2AE1.
CC DR InterPro: IPR002198; ADH_short.
CC DR Pfam: PF00106; adh_short.1.
CC DR PRINTS: PR00080; SDRFAMILY.
CC DR PROSITE: PS00661; ADH_SHORT.1.
CC KM Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
CC FT NP_BIND 11 35 NADP (By SIMILARITY).
CC FT ACT_SITE 156 156 BY SIMILARITY.
CC SQ SEQUENCE 247 AA; 25724 MW; 91EBF9409C777FE20 CRC64;

Query Match 13.2%; Score 215; DB 1; Length 247;
Best Local Similarity 27.5%; Pred. No. 1.1e-10;
Matches 75; Conservative 44; Mismatches 112; Indels 42; Gaps 10;

QY 5 MNGGVCVVTGASRGIGRIALQLCAGATVYITGRHDLTR-VVAQENAGSLGGCCVPPV 63
DB 4 LTAQVALVTGASRGIGRIALTAATGKRVVNVASSTADAVAAETIANGGEIAVOA 63
QY 64 DSSQSEVETLFEQYDREOQGRDLVYNNAYAGVOTILTRNKAFETPASMDDINN 123
DB 64 NVANADEVDQIL-KTTLDKTSRIDVLYNN--AGI-----TRDTLLIRKLEDMQAVIDLN 115
QY 124 LRGHFCVSYGARLWVPAGOGILIVISS-----PGLQYMNVPYGVGAACDKLAAD 176
DB 116 ILGVFLCTKRAVSKMLKOKSGRIINITSVAGMGMPGQAN-----YAAKAGVIGFTKT 169
QY 177 CAHELRRHGVSCVSLMPGVOTELKEHMAKEVLODPVLKQFSAFSATTELSGCV 236
DB 170 VAKELASRGVTVNAAPGIATD-MTENLMAEPILQFIPLARY-----GGPEEVAG-TI 221
QY 237 VALATDPNILSLSGKVLPSCLARRYGLRDV 269
DB 222 RFLATDPAAYITGQTF-----NDG 242

RESULT 9
TRN2_HYONI
ID TRN2_HYONI STANDARD; PRT: 260 AA.
AC P50164;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tropilone reductase-II (EC 1.1.1.236) (TR-II).
GN TR2.
OS Hyoscyamus niger (Henbane).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Solanales; Solanaceae; Hyoscyamus.
OX NCBI_Taxid=4079;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue-Cultured Root;

```

CC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
CC eucosids II; Myrtales; Lythraceae; Cuphea.
CC NCBI_TaxID=3930;
CC [1]
CC SEQUENCE FROM N.A.
CC RA MEDLINE=92293104; PubMed=1376402;
CC Klein B., Pawlowski K., Hoerlcke-Grandpierre C., Schell J.,
CC Toepfer R.;
CC RA Isolation and characterization of a cDNA from *Cuphea lanceolata*
CC RT encoding a beta-ketoacyl-ACP reductase.";
CC Mol. Gen. Genet. 233:122-128(1992).
CC -1 CARBONYL ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
CC NADP(+) -> 3-oxoacyl-[acyl-carrier protein] + NADPH.
CC -1 PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
CC PATHWAY.
CC -1 SUBUNIT: HOMOTETRAMER (PROBABLE).
CC -1 SUBCELLULAR LOCATION: CHLOROPLAST AND NON-PHOTOSYNTHETIC
CC PLASTIDS.
CC -1 SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X64566; CAA45866.1; -;
CC DR PIR, S19832; S19832.
CC DR HSSP, P50162; 1A81.
CC DR InterPro, IPR002198; ADH_short.
CC DR RefSeq, NP_000000000.1; ADH_short.

[illegible]

YCPL_BRAJA STANDARD; PRT: 275 AA.

ID YCPL_BRAJA
AC 045219;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Probable short-chain type dehydrogenase/reductase (EC 1.-.-.-).
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-USDA 110;
RA Tully R.E., Keister D.L.;
RT Cloning and mutagenesis of a cytochrome P-450 locus from
RT Bradyrhizobium japonicum that is expressed anaerobically and
RT symbolically.
RL Appl. Environ. Microbiol. 59:4136-4142(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-USDA 110;
RA MEDLINE=98322110; PubMed=9655913;
RA Tully R.E., van Berkum P., Lovins K.W., Keister D.L.;
RT Identification and sequencing of a cytochrome P450 gene cluster from
RT Bradyrhizobium japonicum.
RL Biochim. Biophys. Acta 1998:243-255(1998).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL: U12678; AAC28892.1; .
DR HSSP: P08074; 1CYD.
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; 1.
DR Hypothetical protein; Oxidoreductase.
KW NP_BIND 10 34 MAD OR NADP (BY SIMILARITY).
FT ACT_SITE 159 159 BY SIMILARITY.
SQ SEQUENCE 275 AA; 28827 MW; 312E7070C404DF86 CRC64;

Query Match 13.2%; Score 213.5; DB 1; Length 275;
Best Local Similarity 31.0%; Pred. No. 1,7e-10;
Matches 65; Conservative 34; Mismatches 100; Indels 11; Gaps 6;

YC 7 GGYCVYTGASRGIGRGIALDLCKAGATVYITGRHDLRLRYVAQNO--SLGGCCVPVCD 64
DB 6 GKAAVYTGAGAGIGKACALIAREGGRVVA--DIDSAALACTAQAIAEAGHALALAI 63
YC 65 SSGSESVRTLFEDVDEOGRDLVLYNNAYAGVOTILTRNKAFWEPASMDIDINNV 124
DB 64 IADAAVAAALFENAEER-HFGVDLVNNASA---MHLPRRALILELELAAYDQTMANT 119
YC 125 RGRHFGSVYCARLMPVAGGGLIVYSS-PSGLQYMFVNPVGVGKACDKLAADCAHELRR 183
DB 120 RGLLLCCROAIPRIAYRGGAIVNMSSCGISGDTALTSTYASRAAMMLSSSLATQYGH 179
YC 184 HGVCVSLMPGIYOTELK--EHMAKEEVL 211
DB 180 AQIRCAVAPGLITERTLRQTHLRHQL 209

Y019_THEME STANDARD; PRT: 256 AA.

ID Y019_THEME
AC 056318;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative oxidoreductase TM0019 (EC 1.-.-.-).
GN TM0019.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MSB8 / DSM 3109;
RX MEDLINE=96125234; PubMed=8550425;
RA Kietzin A., Adams M.;
RT "Molecular and phylogenetic characterization of pyruvate and 2-
RT ketolvalerate ferredoxin oxidoreductases from *Pyrococcus furiosus*
RT and pyruvate ferredoxin oxidoreductase from *Thermotoga maritima*."
RL J. Bacteriol. 178:248-257(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA Hart D.H., Hickey E.K., Peterson J.D., Nelson M.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of *Thermotoga maritima*."
RL Nature 399:323-329(1999).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL: X85171; CAAS9459.1; .
DR EMBL: AE001690; AAC55113.1; .
DR HSSP: P19992; 1HDC.
DR TIGR: TM0019; .
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; 1.
DR Hypothetical protein; Oxidoreductase; NADP; Complete proteome.
KW NP_BIND 9 33 NADP (BY SIMILARITY).
FT ACT_SITE 153 153 BY SIMILARITY.
FT CONFLICT 130 136 RGGCVIT -> TRRSRH (IN REF. 1).
SQ SEQUENCE 256 AA; 28078 MW; D68160B1D7980CB CRC64;

Query Match 12.9%; Score 209; DB 1; Length 256;
Best Local Similarity 30.0%; Pred. No. 3.6e-10;
Matches 75; Conservative 41; Mismatches 118; Indels 16; Gaps 9;

YC 5 MNGCVYTGASRGIGRGIALDLCKAGATVYITGRHDLRLRYVAQ--AQSLGGCCVPV 62
DB 2 LEKAAVYTGAGGIGALIAQLFAENGKVVYA--EIDEAGVEREKLRLRGIDVTFVK 59
YC 63 CDSSESESVRTLFEDVDEOGRDLVLYNNAYAGVOTILTRNKAFWEPASMDIDINNV 122
DB 60 TDVADENSVMKMWRT-VEIYGVDLVNN--AAVMV-----KSIFFRPLEEERYVRV 111
YC 123 GLRGHFGSVYCARLMPVAGGGLIVYSSPSGLQYMFVNPVGVGKACDKLAADCAHEL 181
DB 112 NLGPGYICSRCAEMIRKGGVITININASTRAFOSEPTEPYSASKGLVALTISLAVSL 171

QY 182 RRHGVCSVSLPGIYOT-ELKEHMAKEVLODPVLKQFSAFSSAETTELSCVVALA 240
 DB 172 SRKHIRVVSISPGWIESEMKKSLRKKPDLR-PIDHEOPHAGRVGNPLDIAMLCVF-LA 229
 OY 241 TDPNILLSG 250
 DB 230 DDEKAGFITG 239

RESULT 13
 FABG_ARATH STANDARD; PRT; 319 AA.
 AC P33207;
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 40, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 3-oxoacyl-[acyl-carrier protein] reductase, chloroplast precursor
 DE (EC 1.1.1.100) (3-ketoacyl-acyl carrier protein reductase).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 NC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RX MEDLINE=92246853; PubMed=1575676;
 RA Slabas A.R., Chase D., Nishida I., Murata N., Siedbottom C.,
 RA Safford R., Sheldon P.S., Kekwick R.G.O., Hardie D.G.,
 RA MacIntosh R.W.;
 RT "Molecular cloning of higher-plant 3-oxoacyl-(acyl carrier protein)
 RT reductase. Sequence identities with the node-gene product of the
 RT nitrogen-fixing soil bacterium Rhizobium meliloti.";
 RL Biochem. J. 283:321-326(1992).
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
 CC PATHWAY.
 CC -1- SUBUNIT: HOMOTETRAMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST AND NON-PHOTOSYNTHETIC
 CC PLASTIDS.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X64464; CAA45794.1;
 DR PIR; S22416; S22416.
 DR HSSP; P50162; IAE1.
 DR InterPro; IPR002198; ADH_short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KM Fatty acid biosynthesis; Oxidoreductase; NADP; Chloroplast;
 KM Transl. peptide.
 FT TRANSIT 1 59 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 60 319 3-oxoacyl-(acyl-carrier protein)
 FT NP_BIND 81 105 NADP (BY SIMILARITY).
 FT ACT_SITE 226 226 NADP (BY SIMILARITY).
 FT SEQUENCE 319 AA; 33621 MW; BR3DB304023BPFQ CRC64;

Query Match 12.7%; Score 206; DB 1; Length 319;
 Best Local Similarity 24.9%; Pred. No. 8.3e-10;
 Matches 70; Conservative 49; Mismatches 110; Indels 52; Gaps 8;

QY 9 VCVATGASRIGRIGIALQLOKAGATYIT-GRHLDLRVVAEOASQSGGCVPYVCDSSQ 67
 DB 78 VVITIGASRIGRIGIALQLOKAGATYIT-GRHLDLRVVAEOASQSGGCVPYVCDSSQ 137
 OY 68 ESEVRLFEQVDRQOGRDLVYNNAYAGVOTILNRKAFWETPASMDINNNGRGH 127
 DB 138 ATVDAMMKTA-LDKWGTIDVYVNN-AGI-----FRDILLIRKQSQMDVIALNLTV 189
 OY 128 YFCSYVGARLMPVAGGLIYVSS-PSLOYMENVPGVKACDKLAACAEHLRHGV 186
 DB 190 FLCTQAAVKKMKKKRRIINISSVGLIGNIQANTAKGVISFSETPARGASRNI 249
 OY 187 SCVSLMPGIVOTELKEHMAKEVLODPVLKQFSAFSSAETTELSCVVALATDPNIL 246
 DB 250 NVNVCPCGFASDM-----TAEIGE----- 269
 OY 247 SLGKVLPSCDLARRGLRDVGRPVODYLSLSSVLSHVG 287
 DB 270 DMEKKILGTIPILGRYKAEEVAG-LVEFALSPASAYTG 308

RESULT 14
 YXBG_BACSU STANDARD; PRT; 262 AA.
 AC P46331;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical oxidoreductase yxdG (EC 1.1.1.1).
 GN YXBG OR E3BR.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 NC NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / BGSC1A1;
 RX MEDLINE=96093926; PubMed=7584049;
 RA Yoshida K.-I., Seki S., Fujimura M., Miya Y., Fujita Y.;
 RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis
 RT genome between the gnt and lol operons.";
 RL DNA Res. 2:61-69(1995).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB005554; BAA21601.1;
 DR EMBL; Z99124; CAB16020.1;
 DR HSSP; P25529; IFMC.
 DR Subtilist; BG11357; yxdG.
 DR InterPro; IPR002198; ADH_short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KM Hypothetical protein; Oxidoreductase; NAD; Complete proteome.
 FT NP_BIND 10 34 NAD (BY SIMILARITY).
 FT ACT_SITE 155 155 BY SIMILARITY.
 FT SEQUENCE 262 AA; 28232 MW; A7EE37514F07EDFA CRC64;

Query Match 12.6%; Score 204; DB 1; Length 262;
 Best Local Similarity 26.9%; Pred. No. 9.5e-10;
 Matches 61; Conservative 40; Mismatches 116; Indels 10; Gaps 4;

```

OY 5 MNGOVVATGASRGRIGALQLCKAGATVYITGRHLDTLRVAQEAQSLGGCCPVVCD 64
DB 4 LKNTAVITTAAGSIGGATGATFVANEARVITIGINKDMETVDALRNKGQAESEFLD 63
OY 65 SSOSESEVTLFEQVDRBOQRLDVLVNNAYAVQTIINTNRKAWEPASAMDDINNVL 124
DB 64 VSDENSVKAPADGI-KDACGTIDILFNN--AGV---DQEGCAHEVPVDFDRIANDL 116
OY 125 RGHVFCGVARLMPVPGGGLIYVSSPGSLQVFNPNPYCGKACDKLADCAHELRH 184
DB 117 RGFELSKYILPLMLENGSILINTSSMGRADLDKRGYNNAKGIITLTKRAMIDVARN 176
OY 185 GVCVSLMPGIVOTELKEHAKKEEVLDPPYKOFKSAFSSAETTEL 231
DB 177 GIRVNSISPETIETPLIDKLATRE---QEMGEQFRAETNGSRNSDV 220

RESULT 15
HDNA_ECOLI STANDARD; PRT; 255 AA.
AC P25529;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN 7-alpha-hydroxysteroid dehydrogenase (EC 1.1.1.159) (7-alpha-HSDH).
OS HDHA OR HSDH OR B1619 OR Z2624 OR ECS2327.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562, 83334;
RN (1)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-HB101;
RX MEDLINE=91177803; Pubmed=2007545;
RA Yoshimoto T., Higashi H., Kanetani A., Lin X.S., Nagai H., Oyama H.,
RA Kuzano K., Tsuru D.;
RT *Cloning and sequencing of the 7 alpha-hydroxysteroid dehydrogenase
RT gene from Escherichia coli HB101 and characterization of the
RT J. Bacteriol. 173:2173-2179.(1991).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; Pubmed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT *The complete genome sequence of Escherichia coli K-12.*;
RT Science 277:1453-1474(1997).
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97251357; Pubmed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizoduchi K., Mori H., Mori T.,
RA Ochimura K., Nakase S., Nakamura Y., Nishimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampel G., Seki Y., Sivasubram S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Mada C.,
RA Yamamoto Y., Horiuchi T.;
RT *A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.*;
RT DNA Res. 3:363-377(1996).
RN (4)
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; Pubmed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Pofal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grothbeck E.J., Davis N.W., Llin A., Dimalanta E.T., Potamoudis K.,

```

```

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT *Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.*;
RN Nature 409:529-533(2001).
RN (5)
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; Pubmed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
RT *Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.*;
RN DNA Res. 8:11-22(2001).
RN (6)
RP SEQUENCE OF 169-255 FROM N.A.
RA Jefferson R.A.;
RN Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
RN (7)
RP X-RAY CRYSTALLOGRAPHY (2.3 AND 1.8 ANGSTROMS).
RC STRAIN-HB101;
RX MEDLINE=96264882; Pubmed=8672472;
RA Tanaka N., Nonaka T., Tanabe T., Yoshimoto T., Tsuru D., Mitsui Y.;
RT *Crystal structures of the binary and ternary complexes of 7 alpha-
RT hydroxysteroid dehydrogenase from Escherichia coli.*;
RN Biochemistry 35:7715-7730(1996).
CC -1- FUNCTION: 7-ALPHA-DEHYDROXYLATION OF CHOLIC ACID, YIELDING
CC DEOXYCHOLIC ACID AND LITHOCHOLIC ACID, RESPECTIVELY. HIGHEST
CC AFFINITY WITH TAUROCHENOXYCHOLIC ACID.
CC -1- CATALYTIC ACTIVITY: 3-alpha,7-alpha,12-alpha-trihydroxy-5-beta-
CC cholamate + NAD(+) = 3-alpha,12-alpha-dihydroxy-7-oxo-5-beta-
CC cholamate + NADH.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC
CC EMBL: D10497; BAA01384.1;
CC EMBL: AE000257; AAC74691.1;
CC EMBL: D90805; BAA15370.1;
CC EMBL: D90806; BAA15377.1;
CC EMBL: D90808; BAA15407.1;
CC EMBL: AE005385; AAG56608.1;
CC EMBL: AP002358; BAB35750.1;
CC EMBL: M14641; AAA68921.1;
CC PIR: A38527; A38527.
CC PIR: J70951; J70951.
CC PDB: 1AHH; 1A-OCT-96.
CC PDB: 1AHT; 1A-OCT-96.
CC PDB: 1FMC; 08-NOV-96.
CC SWISS-2DPAGE; P25529; COLI.
CC EcoGene; EG10425; hsdA.
CC InterPro: IPR002198; ADH_short.
CC Pfam: PF00106; adh_short; 1.
CC PRINTS: PR00081; GDRHDH.
CC PROSITE: PS00061; ADH_SHORT; 1.
CC Oxidoreductase; NAD; Bile acid catabolism; 3D-structure;
CC Complete proteome.
CC NP_BIND; 18; NAD.
CC ACT_SITE; 146; 146; SUBSTRATE BINDING.
CC ACT_SITE; 159; 159;
CC SEQUENCE; 255 AA; 26778 MW; 66CF70E85B67B6D CRC64;

```

Query Match 12.5%, Score 203.5; DB 1; Length 255;

Best Local Similarity	27.98	Pred. NO. 1e-09;			
Matches	63; Conservative	41; Mismatches	93; Indels	29; Gaps	7;

```
0Y      5 MNGQVCVTTGASRGRIRIALOLCAKATYYITRHRHLDLRYVAOENOSJGGOCVPVYC 64
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     9 LDGCATITGTAGAGTIGKEIATITFTAGASVVASINDANHHVDEILOJGGAFACRC 68
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y     65 SSGSEVTTFEYOVDREGOGRDLVLYNNAVAVOTITLNTENKAFWETPASMDDINNVGL 124
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     69 ITSEDELSTALD - PAISRLKGTLVNNAGSG ----- GPKPF ----- DMPADF 111
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y     125 RGHFCSYYG ---- ARLMVP - --- GGGLIVISSPGSLQYMENV - PYGVKAACDLIA 175
        | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    112 RRAELNVEFPHLSQLVAPMEKNGGCVILITIISMAEKRNKINIMTSYASSKAAASHLYR 171
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y    176 DCAHELRRHGVSQCVSLPMGIYOTLEIKENHAKE --- EVLODPVLKO 218
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    172 NMAFDLGKKNRINGVINGAIPLDALDKSVITPELECKMLDHPTIRR 217
```

Search completed: July 31, 2002, 15:12:57
Job time: 216 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:11:36 ; Search time 13.05 Seconds

(without alignments)
585,840 Million cell updates/sec

Title: US-10-006-163-1

Perfect score: 313
Sequence: 1 MAPMNGQCVVVGASRGIG.....YLPSFLVPMKIIALYTSKF 313Scoring table:
Gapco 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents, AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTOS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	313	100.0	313	2	US-09-019-216-1
2	16	5.1	323	2	US-09-019-216-3
3	10	3.2	244	1	US-08-375-962B-13
4	10	3.2	244	2	US-08-562-114B-13
5	10	3.2	246	4	US-08-729-594A-13
6	10	3.2	246	3	US-09-238-481-2
7	10	3.2	255	4	US-09-036-987A-21
8	10	3.2	255	4	US-09-370-700-21
9	10	3.2	315	3	US-08-793-035-9
10	10	3.2	315	3	US-08-793-035-10
11	9	2.9	256	1	US-08-594-808B-7
12	8	2.6	277	1	US-08-762-129-5
13	8	2.6	526	2	US-08-852-401-3
14	7	2.2	186	4	US-08-858-207A-270
15	7	2.2	243	4	US-09-239-052-2
16	7	2.2	263	6	5229279-4
17	7	2.2	273	1	US-08-252-995D-10
18	7	2.2	273	2	US-08-834-108-10
19	7	2.2	273	6	5512669-4
20	7	2.2	294	4	US-08-973-334-2
21	7	2.2	294	4	US-09-563-869A-2
22	7	2.2	294	4	US-08-549-489-2
23	7	2.2	327	1	US-08-240-049B-13
24	7	2.2	327	1	US-08-259-148A-15
25	7	2.2	327	1	US-08-484-054-15
26	7	2.2	327	1	US-07-876-941A-15
27	7	2.2	327	4	US-08-542-634-17

28	7	2.2	327	4	US-08-477-292-17	Sequence 17, Appl
29	7	2.2	327	5	PCT-US95-13703-17	Sequence 17, Appl
30	7	2.2	371	3	US-09-043-627-10	Sequence 10, Appl
31	7	2.2	416	1	US-08-252-995D-2	Sequence 2, Appl1
32	7	2.2	416	2	US-08-834-108-2	Sequence 2, Appl1
33	7	2.2	436	1	US-08-259-148A-17	Sequence 17, Appl
34	7	2.2	436	1	US-08-484-054-17	Sequence 17, Appl
35	7	2.2	436	2	US-07-876-941A-17	Sequence 17, Appl
36	7	2.2	464	1	US-08-252-995D-6	Sequence 6, Appl1
37	7	2.2	464	2	US-08-834-108-6	Sequence 6, Appl1
38	7	2.2	525	4	US-08-542-634-27	Sequence 27, Appl
39	7	2.2	525	5	PCT-US95-13703-27	Sequence 27, Appl
40	7	2.2	532	4	US-09-008-271A-12	Sequence 12, Appl
41	7	2.2	534	3	US-08-875-223-8	Sequence 8, Appl1
42	7	2.2	540	4	US-08-542-634-25	Sequence 25, Appl
43	7	2.2	540	5	PCT-US95-13703-25	Sequence 25, Appl
44	7	2.2	549	4	US-08-542-634-15	Sequence 15, Appl
45	7	2.2	549	4	US-08-477-292-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-019-216-1
; Sequence 1, Application US/09019216
; Patent No. 5928925
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN SHORT CHAIN DEHYDROGENASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,216
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0475 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT01
; CLONE: 356351
; US-09-019-216-1

Query Match 100.0%; Score 313; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 7, 6e-292;

Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPNAGVCTVGTASRGIGRGIALQLCKAGATVYITGRHLDTLVVAQEAQSLGGQCP 60
DB 1 MAAPNAGVCTVGTASRGIGRGIALQLCKAGATVYITGRHLDTLVVAQEAQSLGGQCP 60
QY 61 VVCDSSQSESEVTELEQVDRQGRDLVNNAYAGVOTIINTFRKAKWETPASMWDIN 120
DB 61 VVCDSSQSESEVTELEQVDRQGRDLVNNAYAGVOTIINTFRKAKWETPASMWDIN 120
QY 121 NVGLGHYFCVYGARLWVPAGOGILVYISSPSLOYEMFNVPYGVGKACDKLADCAHE 180
DB 121 NVGLGHYFCVYGARLWVPAGOGILVYISSPSLOYEMFNVPYGVGKACDKLADCAHE 180
QY 181 LRRHGVSCVSLMPGIVOTELLKEHMAKEEVLDDPVLKQFSAFSAETTELSGKCVALA 240
DB 181 LRRHGVSCVSLMPGIVOTELLKEHMAKEEVLDDPVLKQFSAFSAETTELSGKCVALA 240
QY 241 TDPNLTLSGKVLPCSDLARRGVLDVGRPVODYLSLSVLSHVSGLGWLASTYLSFLR 300
DB 241 TDPNLTLSGKVLPCSDLARRGVLDVGRPVODYLSLSVLSHVSGLGWLASTYLSFLR 300
QY 301 VPKWIALYTSKF 313
DB 301 VPKWIALYTSKF 313

RESULT 2

US-09-019-216-3

Sequence 3, Application US/09019216

Patent No. 5928923

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

TITLE OF INVENTION: HUMAN SHORT CHAIN DEHYDROGENASE

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Dr.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/019,216

FILING DATE: Filed Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0475 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 323 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 2315796

US-09-019-216-3

Query Match 5.1%; Score 16; DB 2; Length 323;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VTGASRGIGRGIALQL 27
DB 12 VTGASRGIGRGIALQL 27

RESULT 3

US-08-375-962B-13

Sequence 13, Application US/08375962B

Patent No. 5731195

GENERAL INFORMATION:

APPLICANT: SIMON, ANDRAS; HELLMAN, ULF; WERNSTEDT,

TITLE OF INVENTION: Isolated Nucleic Acid Molecule

TITLE OF INVENTION: Which Codes for A 32 kDa Protein Having 11-Cis Retinol

TITLE OF INVENTION: Dehydrogenase Activity, and Which Associates With p63,

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felte & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect (ASCII standard)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/375,962B

FILING DATE: 20-January-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/258,418

FILING DATE: 6-October-1994

ATTORNEY/AGENT INFORMATION:

NAME: Paasqualini, Patricia A.

REGISTRATION NUMBER: 34,894

REFERENCE/DOCKET NUMBER: LUD 5372.1 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 244 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: E.coli 3-oxoacyl[acyl carrier protein]

NAME/KEY: reductase (FABG)

US-08-375-962B-13

Query Match 3.2%; Score 10; DB 1; Length 244;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VTGASRGIGR 21
DB 10 VTGASRGIGR 19

RESULT 4

US-08-562-114B-13

Sequence 13, Application US/08562114B

Patent No. 5972646
GENERAL INFORMATION:
APPLICANT: ERIKSSON ET AL.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE WHICH CODES FOR A
TITLE OF INVENTION: 32 KDA PROTEIN HAVING 11-CIS RETINOL HYDROGENASE
TITLE OF INVENTION: ACTIVITY, AND WHICH ASSOCIATES WITH P63, A
TITLE OF INVENTION: PORTION OF A RETINOL BINDING PROTEIN RECEPTOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect 5.1 and ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/562,114B
FILING DATE: 22-No. 5972646ember-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/375,962
FILING DATE: 20-January-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rohli, Vlnest
REGISTRATION NUMBER: 37,003
REFERENCE/DOCKET NUMBER: LUD 5372.2 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 244 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: E.coli 3-oxoacyl[acyl carrier protein]reductase
NAME/KEY:
US-08-562-114B-13

Query Match 3.2%; Score 10; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VTGASRGIGR 21
DB 10 VTGASRGIGR 19

RESULT 5
US-08-729-594A-13
Sequence 13, Application US/08729594A
Patent No. 6280997
GENERAL INFORMATION:
APPLICANT: Eriksson, Ulf; Simon, Andreas; Romert, Anna
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE WHICH
TITLE OF INVENTION: CODES FOR A 32 KDA PROTEIN HAVING 11-CIS RETINOL DEHYDROGENASE
TITLE OF INVENTION: ACTIVITY, AND WHICH ASSOCIATES WITH P63, A PORTION OF A
TITLE OF INVENTION: RETINOL BINDING PROTEIN RECEPTOR
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,594A
FILING DATE: 11-October-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/562,114
FILING DATE: 22-No. 6280997ember-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/375,962
FILING DATE: 20-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/258,418
FILING DATE: 10-June-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6280997man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5372.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 244 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: E.coli 3-oxoacyl[acyl carrier protein]reductase
NAME/KEY:
US-08-729-594A-13

Query Match 3.2%; Score 10; DB 4; Length 244;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VTGASRGIGR 21
DB 10 VTGASRGIGR 19

RESULT 6
US-09-238-481-2
Sequence 2, Application US/09238481
Patent No. 6110704
GENERAL INFORMATION:
APPLICANT: Huang, Jianzhong
APPLICANT: McDevitt, Damien
TITLE OF INVENTION: Fads2
FILE REFERENCE: GM10192
CURRENT APPLICATION NUMBER: US/09/238,481
CURRENT FILING DATE: 1999-01-28
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 246
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-238-481-2

Query Match 3.2%; Score 10; DB 3; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VTGASRGIGR 21
DB 9 VTGASRGIGR 18

RESULT 7
US-09-036-987A-21
; Sequence 21, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patli J.
; APPLICANT: Turner, Jan R.
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R.
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-036-987A-21

Query Match 3.2%; Score 10; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GRDVLVYVNA 93
Db 94 GRDVLVYVNA 103

RESULT 8
US-09-370-700-21
; Sequence 21, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patli J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; CURRENT FILING DATE: 1999-08-09

EARLIER APPLICATION NUMBER: US 09/36987
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 21
; LENGTH: 255
; TYPE: PR1
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-21

Query Match 3.2%; Score 10; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GRDVLVYVNA 93
Db 94 GRDVLVYVNA 103

RESULT 9
US-08-793-035-9
; Sequence 9, Application US/08793035
; Patent No. 6011201
; GENERAL INFORMATION:
; APPLICANT: Slabas, Antoni R.
; APPLICANT: White, Andrew
; APPLICANT: Chase, Dianne
; APPLICANT: Elborough, Kelvin
; APPLICANT: Penem, Phillip A.
; TITLE OF INVENTION: B-Ketolacyl ACP Reductase Genes From
; TITLE OF INVENTION: Brassica Napus
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,035
; FILING DATE: 28-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9414622.2
; FILING DATE: 20-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB95/01678
; FILING DATE: 17-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kammerer, Patricia A.
; REGISTRATION NUMBER: 29,775
; REFERENCE/DOCKET NUMBER: MOBT:132
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713.787.1400
; TELEFAX: 713.787.1440
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-793-035-9

Query Match 3.2%; Score 10; DB 3; Length 315;
Best Local Similarity 100.0%; Pred. No. 0.098;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 VTGASRGIG 20
Db 76 VTGASRGIG 85

RESULT 10
US-08-793-035-10
Sequence 10, Application US/08793035
Patent No. 6011201
GENERAL INFORMATION:
APPLICANT: Slabas, Antoni R.
APPLICANT: White, Andrew
APPLICANT: Chase, Dianne
APPLICANT: Elborough, Keiran
APPLICANT: Fentem, Phillip A.
TITLE OF INVENTION: B-Ketocycl ACP Reductase Genes From
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,035
FILING DATE: 28-JUL-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9414622.2
FILING DATE: 20-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB95/01678
FILING DATE: 17-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kammerer, Patricia A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: MOBT:132
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713.787.1400
TELEFAX: 713.787.1440
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 315 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-793-035-10

Query Match 3.2%; Score 10; DB 3; Length 315;
Best Local Similarity 100.0%; Pred. No. 0.098;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 VTGASRGIG 20
Db 76 VTGASRGIG 85

RESULT 11
US-08-594-808B-7
Sequence 7, Application US/08594808B
Patent No. 5804423
GENERAL INFORMATION:
APPLICANT: Klasen, Ralf

APPLICANT: Blinger-Meyer, Stephanle
APPLICANT: Sahn, Hermann
APPLICANT: Hollenberg, Cornelies P
TITLE OF INVENTION: MICROBIOLOGICAL METHOD OF MAKING
TITLE OF INVENTION: 5-KETOGLUCONATE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: The Firm of Karl F. Ross, PC
STREET: 5676 Riverdale Ave.
CITY: Bronx
STATE: New York
COUNTRY: USA
ZIP: 10471-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/594,808B
FILING DATE: 07-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Jonathan E
REGISTRATION NUMBER: 26,963
REFERENCE/DOCKET NUMBER: 19893
TELECOMMUNICATION INFORMATION:
TELEPHONE: (718) 884-6600
TELEFAX: 718/601-1099
TELEX: 620428
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-594-808B-7

Query Match 2.9%; Score 9; DB 1; Length 256;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 VTGASRGIG 20
Db 16 VTGASRGIG 24

RESULT 12
US-08-762-129-5
Sequence 5, Application US/08762129
Patent No. 5756299
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: A NOVEL HUMAN CARBONYL REDUCTASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,129
FILING DATE: Herewith
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0171 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 118519
US-08-762-129-5

Query Match 2.6%; Score 8; DB 1; Length 277;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 86 LDVLYNNA 93
|||||||
DB 84 LDVLYNNA 91

RESULT 13
US-08-852-401-3
Sequence 3, Application US/08852401
Patent No. 5976836
GENERAL INFORMATION:
APPLICANT: Weber, J. Mark
APPLICANT: Heesler, Paul E.
APPLICANT: Larsen, Peter E.
APPLICANT: Luu, B. Minh
TITLE OF INVENTION: Methods and Compositions for Enhancing
TITLE OF INVENTION: Erythromycin Production
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
STREET: 2 Prudential Plaza, 180 N. Stetson, Suite
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,401
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mueller, Lisa L.
REGISTRATION NUMBER: 38,978
REFERENCE/DOCKET NUMBER: PER159P0030US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-616-5400
TELEFAX: 312-616-5460
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 526 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-852-401-3

Query Match 2.6%; Score 8; DB 2; Length 526;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 208 EEVLYDPV 215
|||||||
DB 382 EEVLYDPV 389

RESULT 14
US-08-858-207A-270
Sequence 270, Application US/08858207A
Patent No. 6348328
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: No. 6348328el Compounds
NUMBER OF SEQUENCES: 552
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/858,207A
FILING DATE: 09-May-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/017670
FILING DATE: 14-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50475
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 270:
SEQUENCE CHARACTERISTICS:
LENGTH: 186 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6348328e
US-08-858-207A-270

Query Match 2.2%; Score 7; DB 4; Length 186;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 87 DVLVYNN 93
|||||||
DB 83 DVLVYNN 89

RESULT 15

US-09-239-052-2
: Sequence 2, Application US/09239052
: Patent No. 6346395
: GENERAL INFORMATION:
: APPLICANT: Holmes, David J.
: APPLICANT: Zhong, Yiyi
: APPLICANT: Debouck, Christine
: APPLICANT: Jaworski, Deborah D.
: APPLICANT: Wang, Min
: APPLICANT: Warren, Richard L.
: APPLICANT: Kosmatka, Anna L.
: APPLICANT: McDevitt, Damien
: APPLICANT: Ingraham, Karen A.
: APPLICANT: Chalke, Allison F.
: APPLICANT: So, Chi Young
: APPLICANT: Wallis, Nicola G.
: APPLICANT: Pearson, Stewart C.
: TITLE OF INVENTION: FabG
: FILE REFERENCE: GM10191
: CURRENT APPLICATION NUMBER: US/09/239,052
: CURRENT FILING DATE: 1999-01-27
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 243
: TYPE: PRT
: ORGANISM: Streptococcus pneumoniae
US-09-239-052-2

Query Match 2.2%; Score 7; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 87 DVLVNNNA 93
|||
Db 83 DVLVNNNA 89

Search completed: July 31, 2002, 15:13:53
Job time: 137 sec